UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Computational Biology and Complex Systems Research Group

CMCiB Comparative Medicine & Bioimage Centre of Catalonia

IGTP Germans Trias i Pujol Research Institute

Daily report 25-05-2020

Analysis and prediction of COVID-19 for EU-EFTA-UK and other countries

Contact: clara.prats@upc.edu

With the financial support of: MINISTERIO DE CIENCIA E INNOVACIÓN "la Caixa"

Foreword

The present report aims to provide a comprehensive picture of the **pandemic situation of COVID-19** in the EU countries, and to be able to foresee the situation in the next coming days.

We employ an **empirical model**, verified with the evolution of the number of confirmed cases in previous countries where the epidemic is close to conclude, including all provinces of China. The model does not pretend to interpret the causes of the evolution of the cases but to permit the **evaluation of the quality of control measures made in each state** and a **short-term prediction of trends**. Note, however, that the effects of the measures' control that start on a given day are not observed until approximately 7-10 days later.

The model and predictions are based on two parameters that are daily fitted to available data:

- ✓ *a*: the velocity at which spreading specific rate slows down; the higher the value, the better the control.
- ✓ *K*: the final number of expected cumulated cases, which cannot be evaluated at the initial stages because growth is still exponential.

We show an individual report with 8 graphs and a table with the **short-term predictions** for different countries and regions. We are adjusting the model to **countries and regions** with at least 4 days with more than 100 confirmed cases and a current load over 200 cases. The **predicted period** of a country depends on the number of datapoints over this 100 cases threshold, and is of 5 days for those that have reported more than 100 cumulated cases for 10 consecutive days or more. For short-term predictions, we assign higher weight to last 3 points in the fittings, so that changes are rapidly captured by the model. The whole methodology employed in the inform is explained in the last pages of this document.

In addition to the individual reports, the reader will find an initial dashboard with a brief analysis of the situation in EU-EFTA-UK countries, some summary figures and tables as well as **long-term predictions** for some of them, when possible. These long-term predictions are evaluated without different weights to data-points. We also discuss a specific issue every day.

Martí Català
Pere-Joan Cardona, PhD
Comparative Medicine and Bioimage Centre of Catalonia; Institute for Health Science Research Germans Trias i Pujol

Clara Prats, PhD
Sergio Alonso, PhD
Enric Álvarez, PhD
Miquel Marchena
David Conesa
Daniel López, PhD
Computational Biology and Complex Systems; Universitat Politècnica de Catalunya - BarcelonaTech

With the collaboration of: Guillem Álvarez, Oriol Bertomeu, Laura Dot, Lavínia Hriscu, Helena Kirchner, Daniel Molinuevo, Pablo Palacios, Sergi Pradas, David Rovira, Xavier Simó, Tomás Urdiales

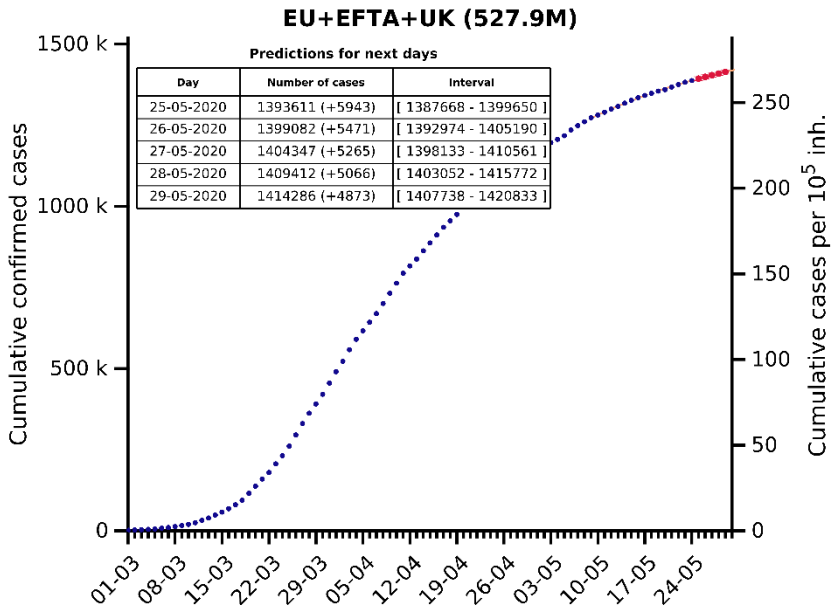
PJC and MC received funding from "la Caixa" Foundation (ID 100010434), under agreement LCF/PR/GN17/50300003; CP, DL, SA, MC, received funding from Ministerio de Ciencia, Innovación y Universidades and FEDER, with the project PGC2018-095456-B-I00;

Disclaimer: These reports have been written by declared authors, who fully assume their content. They are submitted daily to the European Commission, but this body does not necessarily share their analyses, discussions and conclusions.

(0) Executive summary – Dashboard

Global EU+EFTA+UK trends and needs

Global trend in EU+EFTA+UK keeps being positive, with a **ρ_7<1 since beginning April**. We are at the level of 6,000-5,000 daily new cases, and de deaths' curve is tending to saturation as well.



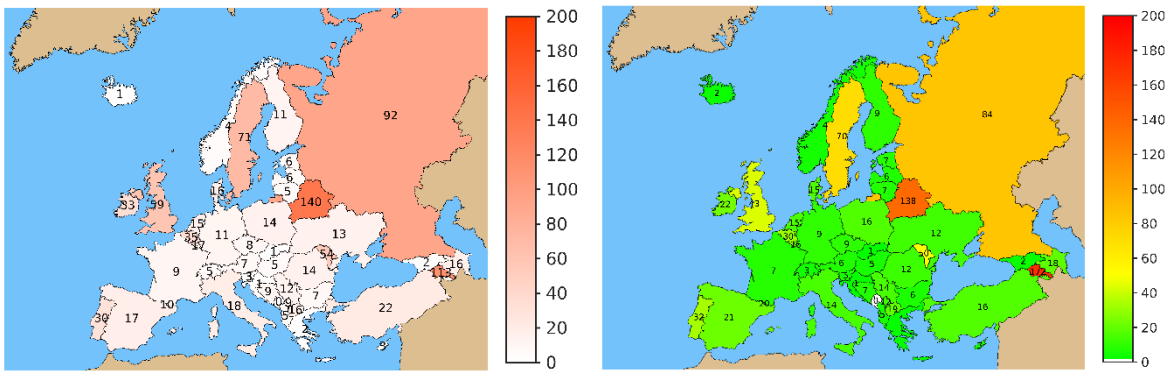
Looking at the bottom of this page, **there are three countries that stand out in the maps: Russia, Belarus, and Armenia**. Russia is the first European country in number of cases (335,882), and the second in the world after the USA. However, the number of new daily cases started to decline about two weeks ago and it is to be expected that they will gradually get closer to the control of the epidemic. They currently have more than 9,000 new cases daily. Belarus is a much smaller country than Russia, it has 35,244 cases, but this corresponds to a higher incidence than its neighbor. At this moment they have not yet started the decline in the number of daily cases. In fact, they are at the top of the mountain chain with an average of about 930 new cases daily. A common feature in Russia and Belarus is low mortality, but it must be said that such low values are not at all consistent with mortality in any other country in the world. Therefore, it clearly indicates that this information is most likely incorrect. Armenia is a small country of only about 3 million people, with currently 6,302 cases. The problem is that its growth is still very important, it does not seem at the moment that even a slowdown has begun.

The analysis is focused on **distinguishing fluctuations in daily new cases**, which are inherent to data, **from possible secondary outbreaks**.

Trends for specific countries

UK is almost to overcome the 260,000 cases, and shows the highest number of reported active cases (40,376). In relative terms, highest 14-day attack rate belongs to **Sweden** (72.5 per 100,000 inhabitants), closelly followed by **UK** (60.8). Other countries seem to maintain the low risk, despite deconfinement measures are being applied.

The map in the left shows current **A_{14}**. The map in the right shows current **EPG**.



Situation and trends per country

Table of current situation in EU countries. Colour scale is relative except when indicated, this means that it is applied independently to each column, and distinguishes best (green) form worst (red) situations according to each of the variables. Last column (**EPG_{EST}**) indicates EPG assessed with **estimated real 14-day attack rate** (see report from 22/04 for details). **EPG_{REP}** is calculated with **data reported by countries**. EPG_{REP} and EPG_{EST} **cannot be compared between them** because scales are different, but can be independently used for estimating risk of countries according to reported or estimated real situation, respectively.

	Reported data								Indexes		
Country	Cumulative cases	Attack rate /10^5 inh.	Cumulative deaths	Mortality /10^5 inh.	Active cases (last 14 days)	14-day attack rate /10^5 inh.	Estimated active cases (last 14 days)	Estimated 14-day attack rate /10^5 inh.	ρ_7 [1]	EPG_{REP} [2]	EPG_{EST} [3]
United Kingdom	259,559	390.7	36,793	55.4	40,376	60.8	596,907	879.3	0.73	44	638
Spain	235,772	508.7	28,752	62.0	8,395	18.1	114,729	244.0	1.38	25	336
Italy	229,858	386.8	32,785	55.2	10,788	18.2	155,383	257.0	0.79	14	204
Germany	178,570	218.0	8,257	10.1	8,995	11.0	42,933	51.2	0.86	9	44
France	144,921	223.9	28,367	43.8	5,858	9.1	114,753	175.8	0.80	7	140
Belgium	57,092	502.7	9,280	81.7	4,011	35.3	67,942	586.2	0.88	31	515
Netherlands	45,236	266.3	5,822	34.3	2,609	15.4	34,568	201.7	0.98	15	197
Sweden	33,459	340.1	3,998	40.6	7,137	72.5	93,956	930.3	0.99	72	918
Switzerland	30,653	357.7	1,641	19.1	431	5.0	2,358	27.2	0.68	3	18
Portugal	30,623	295.2	1,316	12.7	3,042	29.3	13,754	134.9	1.08	32	146
Ireland	24,639	521.3	1,608	34.0	1,643	34.8	11,200	226.8	0.67	23	152
Poland	21,326	55.8	996	2.6	5,330	13.9	28,466	75.2	1.12	16	84
Romania	18,070	91.4	1,179	6.0	2,708	13.7	18,832	97.9	0.84	11	82
Austria	16,439	188.7	640	7.3	652	7.5	2,548	28.3	0.85	6	24
Denmark	11,360	198.9	562	9.8	931	16.3	4,755	82.1	0.93	15	77
Czech Republic	8,957	84.4	315	3.0	834	7.9	3,138	29.3	1.13	9	33
Norway	8,309	154.8	235	4.4	210	3.9	622	11.5	1.03	4	12
Finland	6,579	119.6	307	5.6	617	11.2	3,334	60.2	0.79	9	48
Luxembourg	3,992	693.1	110	19.1	106	18.4	303	48.4	0.93	17	45
Hungary	3,756	38.5	491	5.0	472	4.8	6,555	67.9	0.93	5	63
Greece	2,878	25.7	171	1.5	162	1.4	973	9.3	0.61	1	6
Bulgaria	2,433	34.1	130	1.8	468	6.6	2,709	39.0	0.83	5	32
Croatia	2,244	53.3	99	2.3	57	1.4	NA	NA	0.93	1	NA
Estonia	1,823	138.9	64	4.9	84	6.4	NA	NA	1.17	7	NA
Iceland	1,804	495.3	10	2.7	3	0.8	NA	NA	0.71	1	NA
Lithuania	1,623	55.8	63	2.2	144	5.0	NA	NA	1.26	6	NA
Slovakia	1,509	27.7	28	0.5	52	1.0	NA	NA	0.90	1	NA
Slovenia	1,509	72.6	107	5.1	52	2.5	382	18.4	2.45	6	45
Latvia	1,047	53.1	22	1.1	108	5.5	NA	NA	1.08	6	NA
Cyprus	935	79.9	17	1.5	37	3.2	NA	NA	1.13	4	NA
Malta	610	142.2	6	1.4	114	26.6	NA	NA	NA	NA	NA
Liechtenstein	83	215.3	1	2.6	0	0.0	NA	NA	NA	NA	NA
	Scale										
	Worst	Worst	Worst	Worst	Worst	Worst	Worst	Worst	2.0	100	1000
	Best	Best	Best	Best	Best	Best	Best	Best	0.0	0	0

[1] ρ_3 is the average of 7 consecutive ρ, but can still fluctuate. [2,3] EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

Highlights for countries with highest number of reported cases

- ✓ UK seems to be consolidating the decreasing trend, with an $\rho_7<1$ for a couple of weeks.
- ✓ Spain's ρ_7 is affected by two one-day increases in new cases. Globally, its trend is positive.
- ✓ France, Italy and Germany follow expected trends, all of them in the control phase.

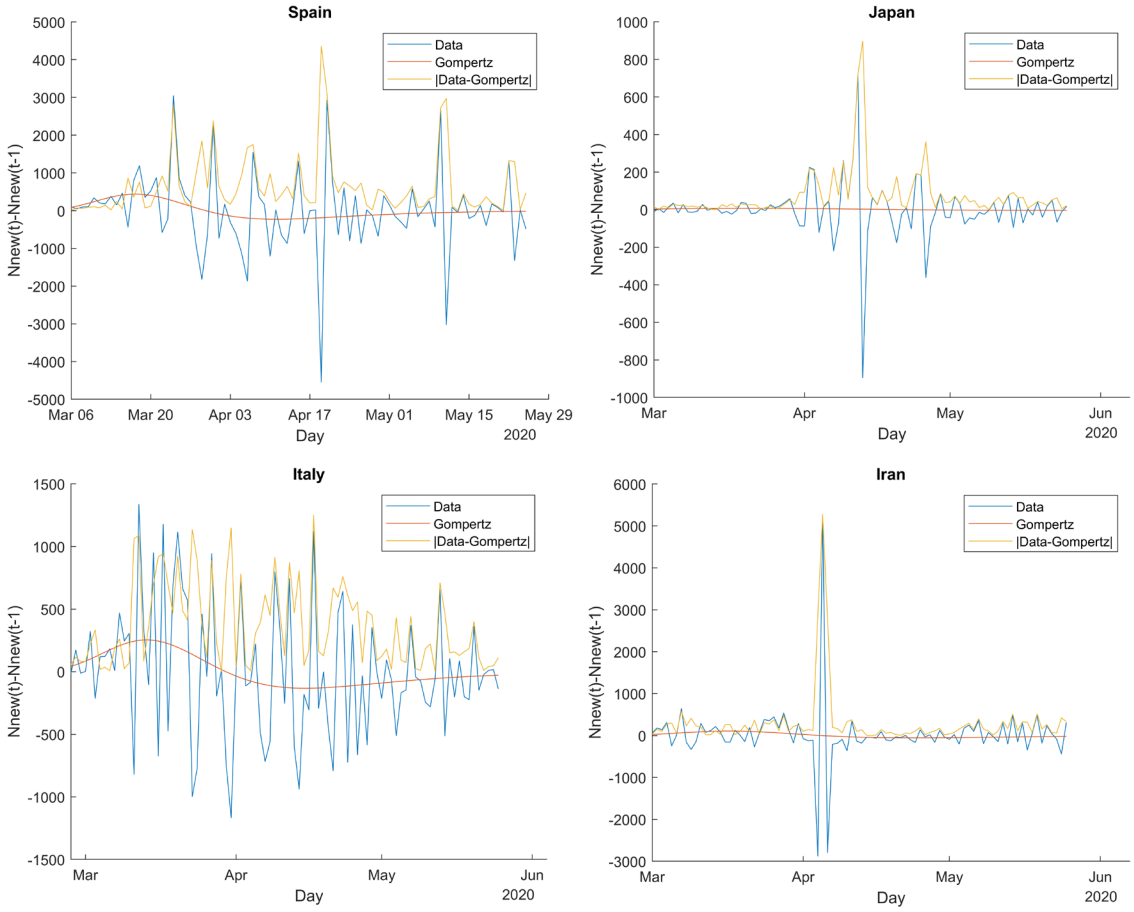
Analysis: Fluctuations and risk of subsequent outbreaks.

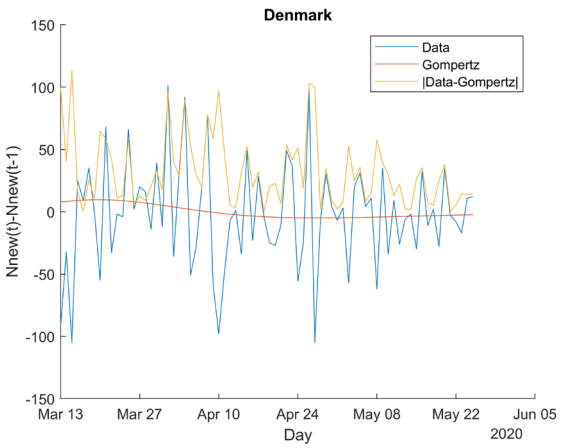
The situation in Europe, a part of some important exceptions like United Kingdom and Sweden, is actually evolving towards a controlled state. The **number of new cases in the great majority of the countries is decreasing and the health systems are recovering their usual conditions**. Under such situation the main concern of the authorities is the **risk of new outbreaks** and therefore, it is important to detect as soon as possible such abnormal increase in new cases. However, the inspection of the evolution of the new cases in all the countries show strong deviations with the averaged behavior. Such deviations are twofold, first related to the intrinsic dynamics of the stochastic nature of the infection process and second because of all the external constrains on the correct and precise reporting of the data.

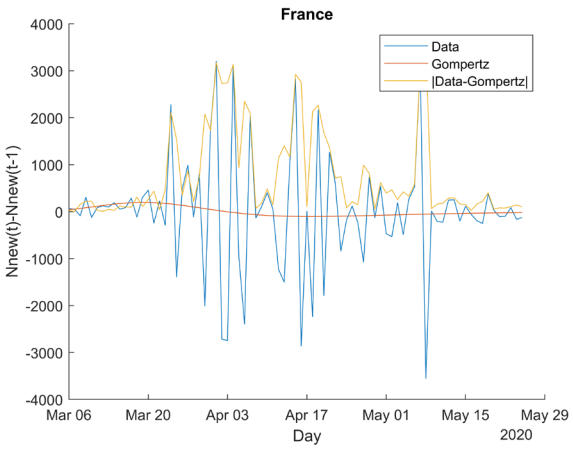
A possible outbreak may be masked inside these fluctuations of the own measurements. Therefore, we have to identify first the magnitude of these fluctuations inside the series of data to be able to define thresholds to the appearance of new outbreaks.

Fluctuations on the new cases

A first inspection of the data shows that the series are noisy. The series can be described by an average value with some fluctuations, $N_{new}(t) = N_{new,0} + \xi(t)$, where $\xi(t)$ is assumed to follow a stochastic Gaussian distribution with a variance σ_1. In order to avoid any assumption about the averaged behavior $N_{new,0}$ and bias the results we consider the fluctuations as $N_{new}(t) - N_{new}(t-1)$. If $N_{new,0}$ does not change to much, we obtain a second stochastic magnitude $\eta(t) = \xi(t) - \xi(t-1)$, which also follows an stochastic Gaussian distribution with variance $\sigma_2^2 = 2\,\sigma_1^2$. The evolution of this quantity is shown in the next figure for six different countries (blue curves). The set of values are renormalized with the Gompertz curve in order to avoid the drift of the values due to the deterministic change of the quantity $N_{new,0}(t) - N_{new,0}(t-1)$ (yellow curves).



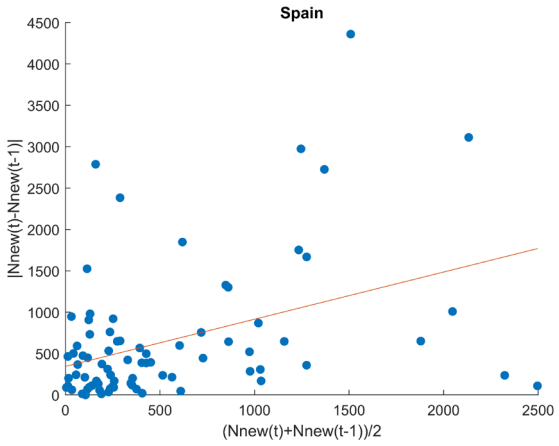


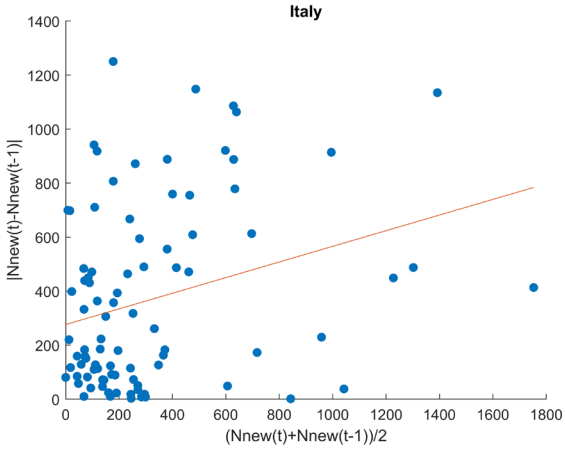


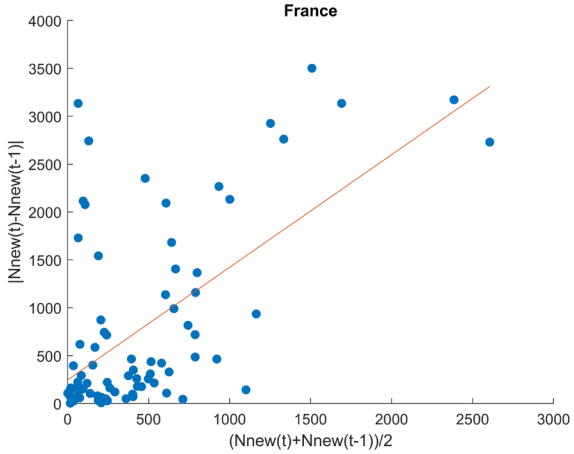
We observe that there is a significant increase of the fluctuations in the middle of the graphics, corresponding with the larger values of the new dates $N_{new,0}(t)$.

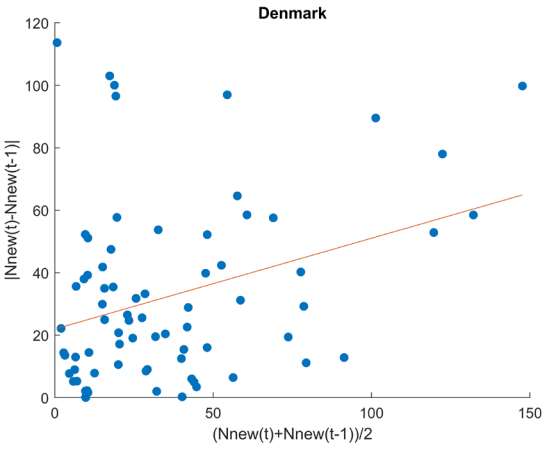
Relative fluctuations on the new cases

The fluctuations depend on the current number of new cases. Fluctuations are smaller at the initial states than in the intermediate situation around the maximum of new cases. In the next figure we analyze the series of the **fluctuations in front of the average value of the new cases**. We see that the amplitude of the fluctuations increases with the actual value of the number of new cases $N_{new}(t)$ averaged for the 2 consecutive days (t and t-1).





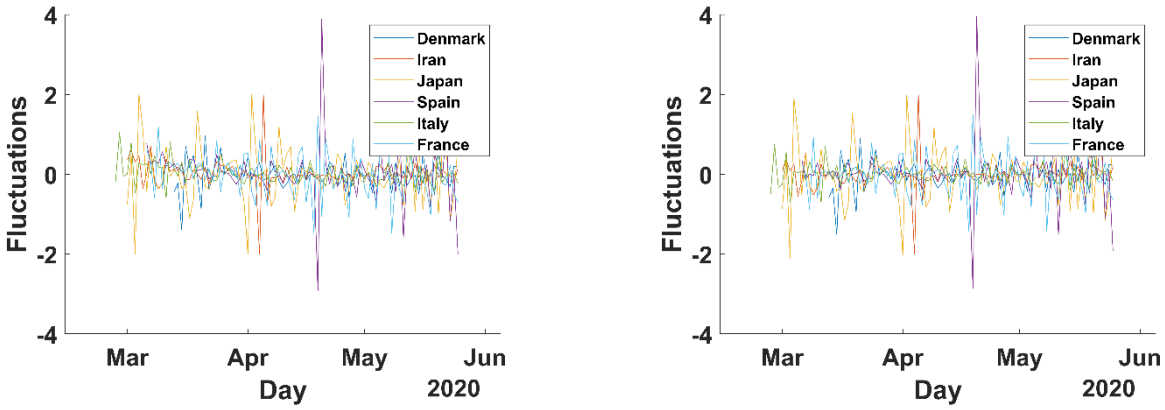




The lines in the figure permit to show the increase of the fluctuations with the number of new cases. To obtain a more reliable quantity we consider the relative error defined as:

$$Fluctuations = \frac{2 \cdot \left(N_{new}(t) - N_{new}(t-1)\right)}{\left(N_{new}(t) + N_{new}(t-1)\right)}$$

trying to homogenize the fluctuations along the time. The evolution of these fluctuations is shown in the third figure for six different countries. In general, the dispersion of the relative fluctuations remains relatively constant along the time in the six countries (figure in the left).



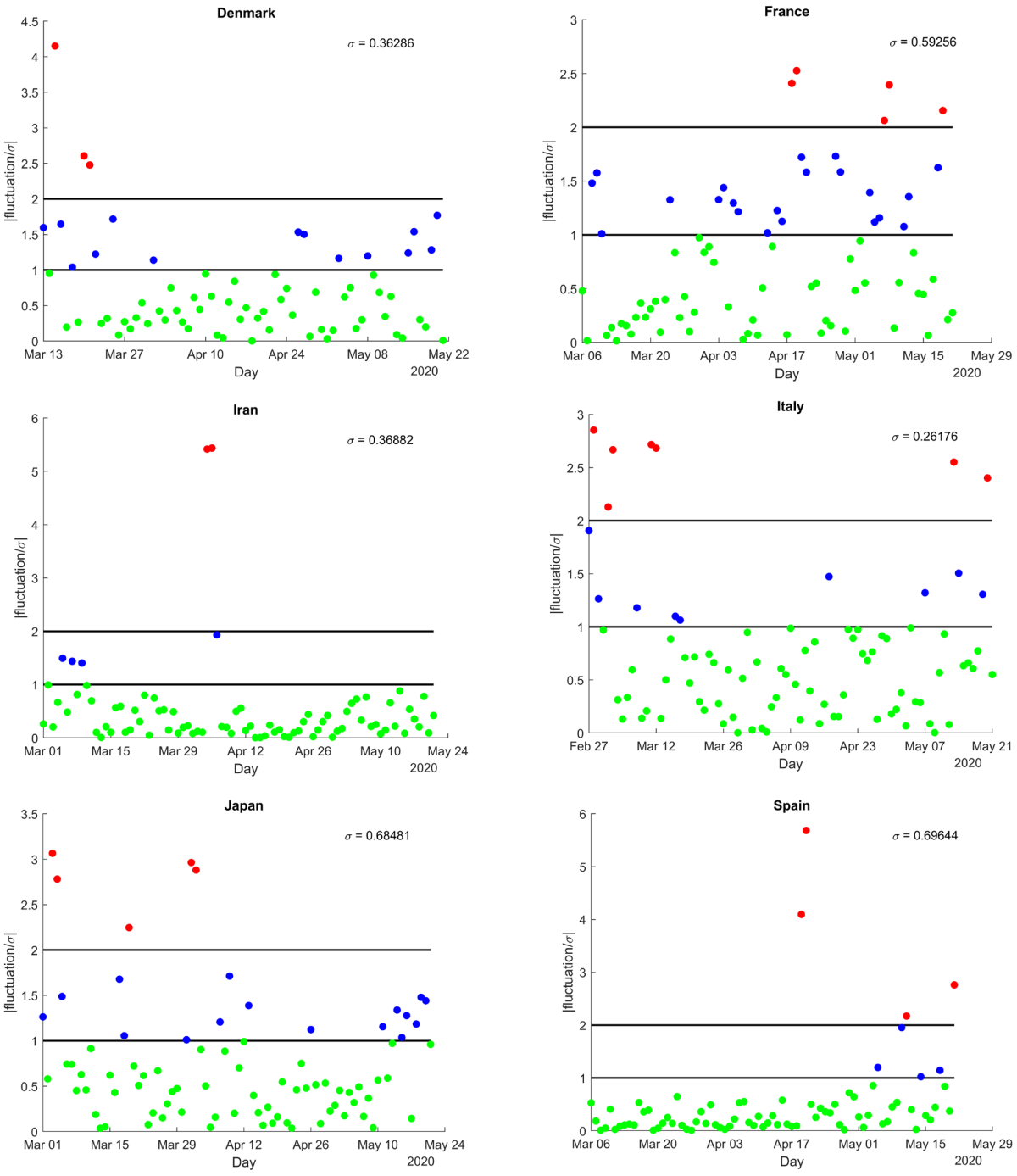
There is however a net increase at the initial of the epidemics because the strong exponential growth (therefore, the magnitude $N_{new,0}(t) - N_{new,0}(t-1)$ is clearly different than zero). We employ the Gompertz model to evaluate this quantity and to renormalize the corresponding curve, see the second panel where such initial increase disappears (figure on the right). In this case, the quantity *Fluctuations* is zero mean and has a variance of sigma approximately equal to $\sigma_2 \cdot N_{new}(t)$.

Therefore, in the previous figure we have the series of fluctuations for each country. We have the possibility to calculate the variance of these fluctuations for each country and **characterize the typical magnitude of their fluctuations giving the values of the daily new cases**.

Twice the variance of relative fluctuations as threshold for new outbreaks

We analyze the temporal evolution of the fluctuations in comparison of the typical variances appearing in the daily new cases to evaluate the possible appearance of new outbreaks or big disturbances in the protocol in the reporting of the daily new data, which is probably the biggest inconvenience to check for outbreaks.

Next, we show the **fluctuations obtained during the epidemic for the six different countries**, depending on the country the level of the fluctuations is different and therefore the definition of outbreak has to adapt to the corresponding level of the country under consideration. **If the new data of a particular day for the next weeks cross the 2σ level (red points in the figures) the authorities may have to be alert for the next days' evolution**.



Note that the biggest deviations of the data appear in pairs. This is because they are **typically associated to the update of the new cases recollected and not previously reported for any reason**. Probably, for a better estimation one may have to remove such pairs of points which appear with opposite signs and calculate again the variance.

In conclusion, **for a good estimation of the beginning of an outbreak it may be useful to have a threshold value given by the typical fluctuations of the number of new cases to detect possible candidates to outbreaks with enough anticipation**.

Situation and trends in other countries

Table of current situation in a sample of non-EU countries. Colour scale is relative except when indicated, this means that it is applied independently to each column, and distinguishes best (green) form worst (red) situations according to each of the variables. EPG_{REP} and EPG_{EST} **cannot be compared between them** because scales are different, but can be independently used for estimating risk of countries according to reported or estimated real situation, respectively.

	Reported data								Indexes		
Country	Cumulative cases	Attack rate /10^5 inh.	Cumulative deaths	Mortality /10^5 inh.	Active cases (last 14 days)	14-day attack rate /10^5 inh.	Estimated active cases (last 14 days)	Estimated 14-day attack rate /10^5 inh.	ρ_7 [1]	EPG_{REP} [2]	EPG_{EST} [3]
United States of America	1,643,238	496.4	97,720	29.5	313,439	94.7	2,036,341	615.2	0.99	94	610
Brazil	363,211	170.9	22,666	10.7	200,512	94.3	1,562,490	735.1	1.35	128	995
Russia	344,481	236.1	3,541	2.4	134,793	92.4	NA	NA	0.91	84	NA
Argentina	12,063	26.7	452	1.0	6,287	13.9	37,622	83.2	1.75	24	145
Belarus	36,198	383.1	199	2.1	13,225	140.0	NA	NA	0.99	138	NA
Peru	119,959	363.8	3,456	10.5	52,652	159.7	175,768	533.1	1.05	168	560
Canada	84,688	224.4	6,424	17.0	15,840	42.0	143,142	379.3	0.97	41	366
Mexico	68,620	53.2	7,394	5.7	33,598	26.1	481,154	373.2	1.22	32	456
Chile	69,102	361.5	718	3.8	40,236	210.5	95,608	500.1	1.47	309	735
Saudi Arabia	72,560	208.4	390	1.1	33,512	96.3	NA	NA	1.14	109	NA
Pakistan	56,349	25.5	1,167	0.5	25,408	11.5	61,699	27.9	1.23	14	34
Qatar	43,714	1,517.3	23	0.8	21,194	735.6	NA	NA	1.08	796	NA
India	138,845	10.3	4,021	0.3	71,693	5.3	256,567	19.0	1.34	7	25
Iran	135,701	161.6	7,417	8.8	28,098	33.5	163,671	194.9	1.15	39	225
	Scale										
	Worst	Worst	Worst	Worst	Worst	Worst	Worst	Worst	2.0	100	1000
	Best	Best	Best	Best	Best	Best	Best	Best	0.0	0	0

[1] ρ_7 is the average of 7 consecutive ρ, but can still fluctuate. [2,3] EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

Disclaimer: estimated active cases and estimated 14-day attack rate are assessed by assuming a lethality of 1 % (see report from 20 to 24 April, #37-41). This value can change in countries where suspicious deaths are reported as well (real values would be lower) and in countries where incidence among elderly people was minor (real values would be higher).

Time indicators by country

This tables summarize a few time indicators for each country: time since 50 cases were reported, time interval between an attack rate of $1/10^5$ inhabitants and an attack rate of $10/10^5$ inhabitants, and time interval between attack rates of 10 to 100 per 10^5 inhabitants (only for countries that have overtaken this threshold).

EU+EFTA+UK countries

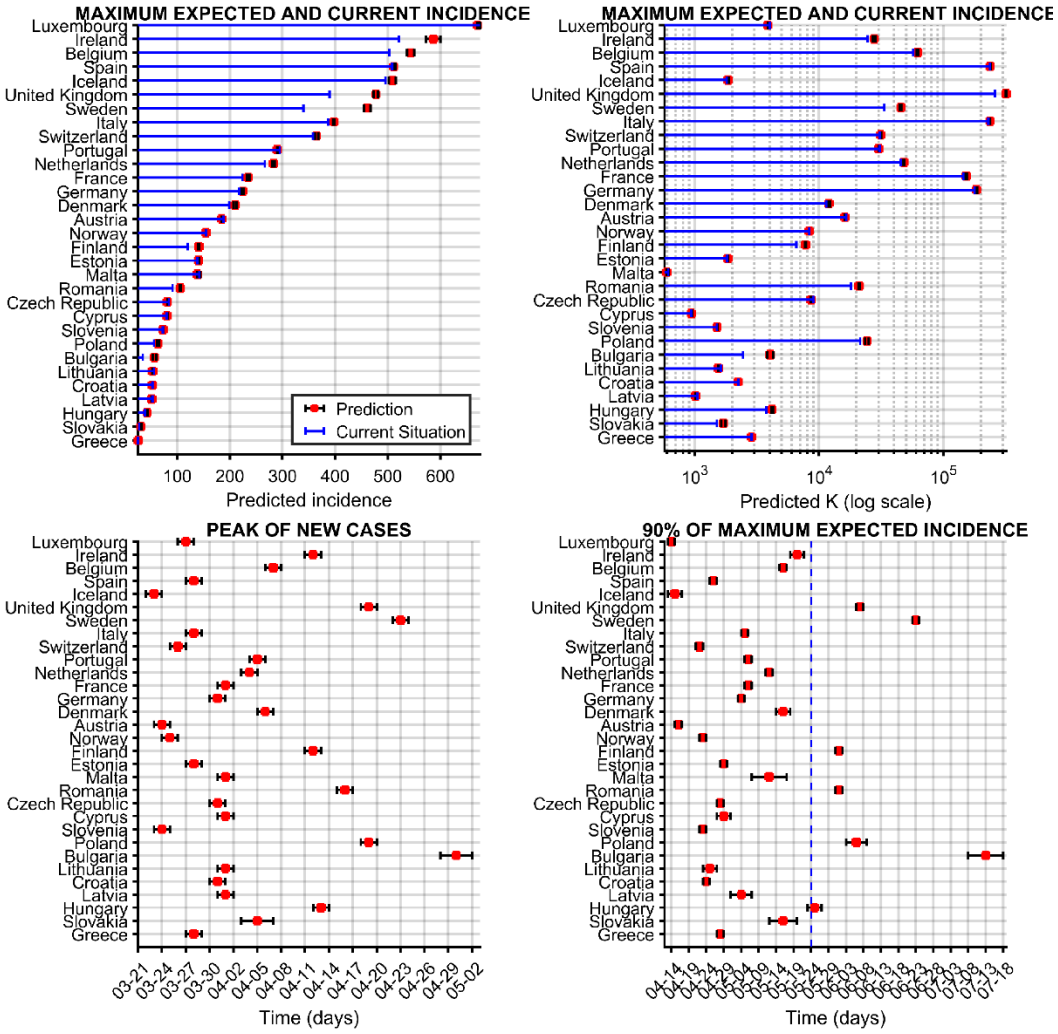
Countries	Days since the first 100 cases	Time interval between 1 and 10 cases / 10^5 inh. (days)	Time interval between 10 and 100 cases / 10^5 inh. (days)
Italy	92	11	16
Germany	86	12	17
France	85	10	20
Spain	85	8	12
United Kingdom	81	10	12
Belgium	80	11	14
Netherlands	80	11	20
Sweden	80	10	28
Norway	80	2	7
Switzerland	80	8	11
Austria	78	10	14
Denmark	77	4	30
Czech Republic	74	11	NA
Finland	74	12	46
Greece	74	18	NA
Iceland	74	5	15
Portugal	73	9	15
Slovenia	73	6	NA
Estonia	72	5	30
Ireland	72	8	18
Poland	72	17	NA
Romania	72	15	NA
Luxembourg	69	6	7
Slovakia	68	24	NA
Bulgaria	67	30	NA
Croatia	67	12	NA
Hungary	66	20	NA
Latvia	66	12	NA
Lithuania	65	9	NA
Malta	64	9	35
Cyprus	63	12	NA

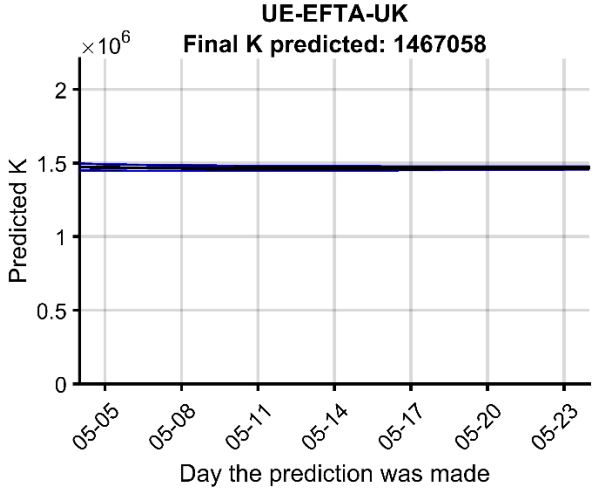
Other countries

Countries	Days since the first 100 cases	Time interval between 1 and 10 cases / 10^5 inh. (days)	Time interval between 10 and 100 cases / 10^5 inh. (days)
Iran	89	11	42
United States of America	84	8	15
Canada	75	11	27
Qatar	75	3	31
Brazil	72	20	34
Saudi Arabia	71	21	29
Chile	70	13	36
Pakistan	70	35	NA
India	70	NA	NA
Russia	69	15	24
Peru	69	18	22
Mexico	68	25	NA
Argentina	67	39	NA
Belarus	56	10	18

Long-term predictions

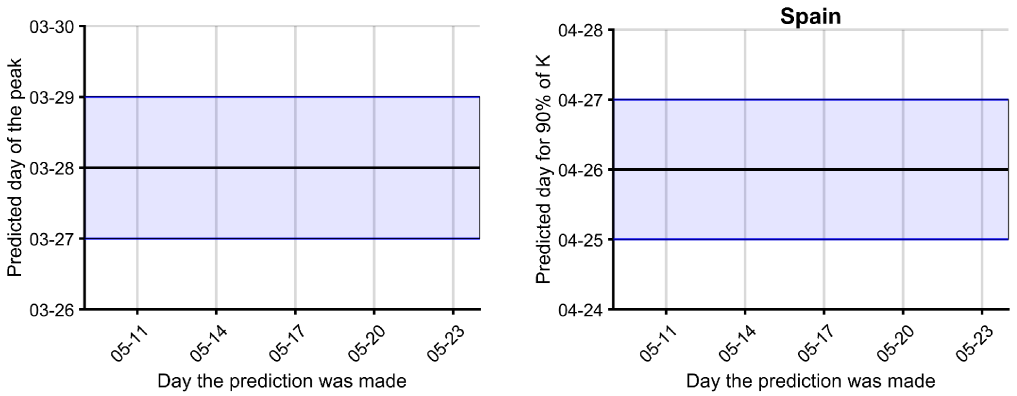
Evaluated with the **whole historical series**. See figure in the next page. Up-left: Predictions of maximum incidences per country (total final expected attack rate per 10^5 inh.). Up-right: Predictions of maximum absolute number of cases per country (K, in log scale). Blue lines indicate current situation. Bottom-left: Time in which peak in new cases was achieved / will be achieved. Bottom-right: Time at which 90 % of K was achieved / will be achieved. Blue dotted line indicates current date. At the end, predicted K for whole EU+EFTA+UK.





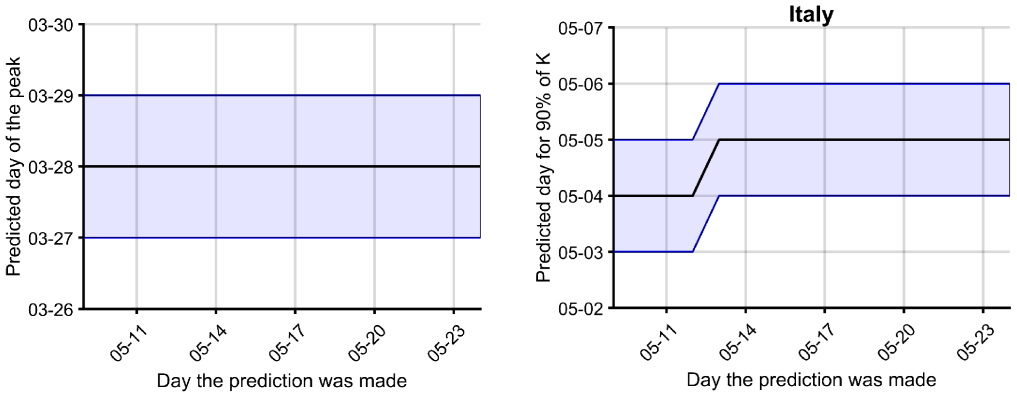
2020-05-24

Predicted day of the peak
03-30
03-29
03-28
03-27
03-26
05-11
05-14
05-17
05-20
05-23
Day the prediction was made


Spain
Predicted day for 90% of K
04-28
04-27
04-26
04-25
04-24
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

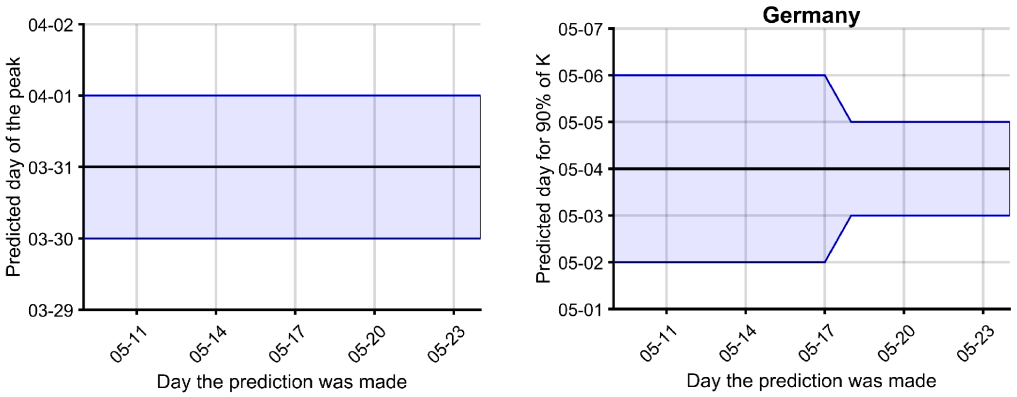
×10^5
Predicted K
3.5
3
2.5
2
1.5
1
0.5
0
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

Predicted day of the peak
03-30
03-29
03-28
03-27
03-26
05-11
05-14
05-17
05-20
05-23
Day the prediction was made


Italy
Predicted day for 90% of K
05-07
05-06
05-05
05-04
05-03
05-02
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

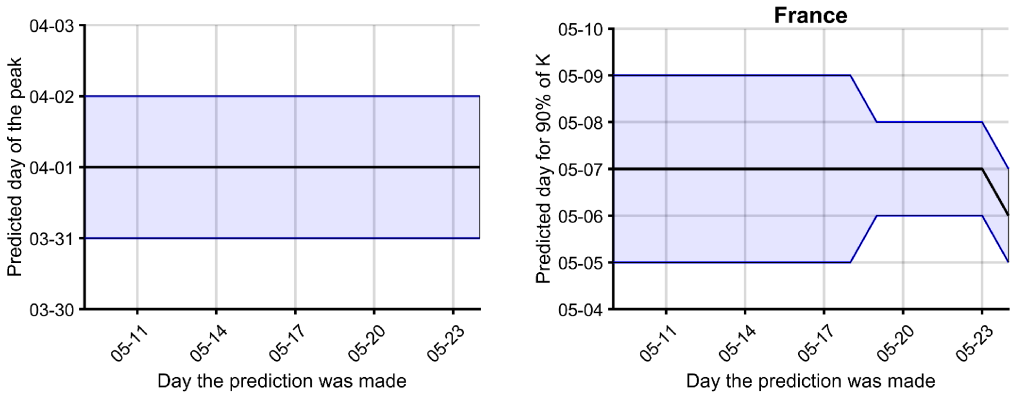
×10^5
Predicted K
3.5
3
2.5
2
1.5
1
0.5
0
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

Predicted day of the peak
04-02
04-01
03-31
03-30
03-29
05-11
05-14
05-17
05-20
05-23
Day the prediction was made


Germany
Predicted day for 90% of K
05-07
05-06
05-05
05-04
05-03
05-02
05-01
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

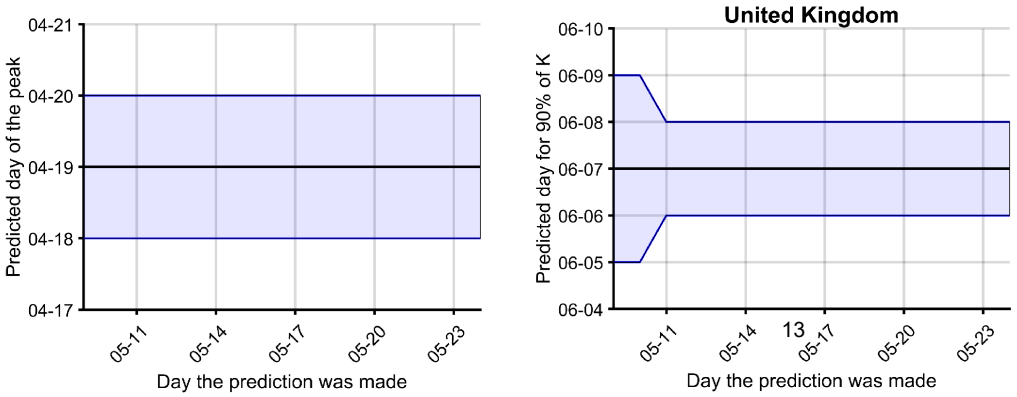
×10^5
Predicted K
2.5
2
1.5
1
0.5
0
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

Predicted day of the peak
04-03
04-02
04-01
03-31
03-30
05-11
05-14
05-17
05-20
05-23
Day the prediction was made


France
Predicted day for 90% of K
05-10
05-09
05-08
05-07
05-06
05-05
05-04
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

×10^5
Predicted K
2
1.5
1
0.5
0
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

Predicted day of the peak
04-21
04-20
04-19
04-18
04-17
05-11
05-14
05-17
05-20
05-23
Day the prediction was made


United Kingdom
Predicted day for 90% of K
06-10
06-09
06-08
06-07
06-06
06-05
06-04
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

×10^5
Predicted K
4
3
2
1
0
05-11
05-14
05-17
05-20
05-23
Day the prediction was made

Situation, trends and long-term predictions in Italian regions[1]

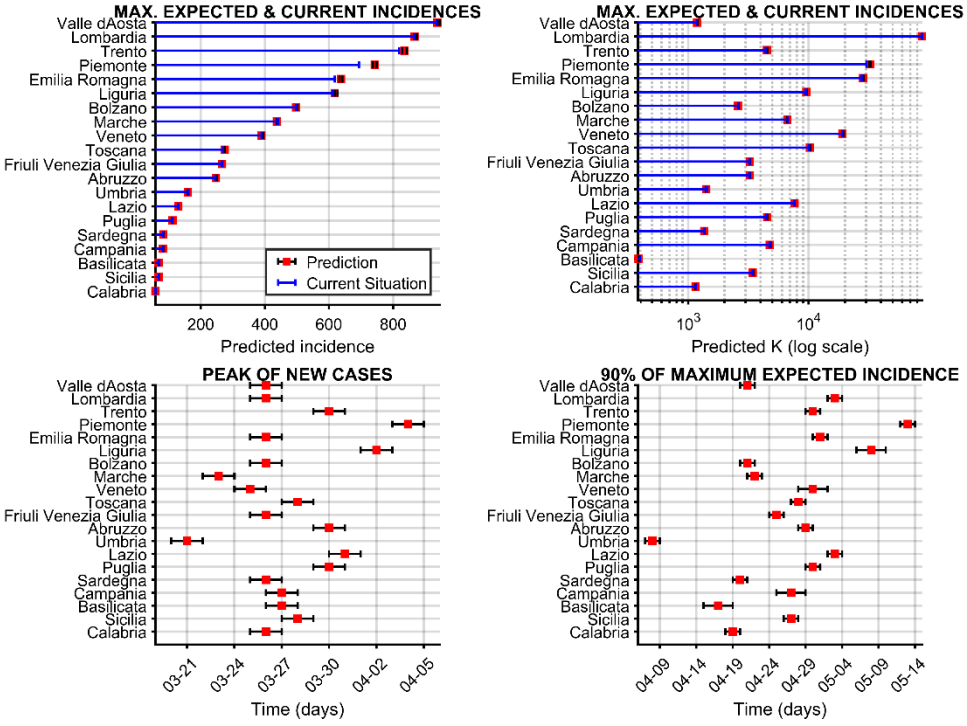
Situation and trends

	Reported data								Indexes		
Country	Cumulative cases	Attack rate /10^5 inh.	Cumulative deaths	Mortality /10^5 inh.	Active cases (last 14 days)	14-day attack rate /10^5 inh.	Estimated active cases (last 14 days)	Estimated 14-day attack rate /10^5 inh.	ρ_7 (1)	EPG_{REP} (2)	EPG_{EST} (3)
Lombardia	87.258	868,9	15.874	158,1	5.387	53,6	95.769	951,9	0,88	47	841
Piemonte	30.228	693,9	3.798	87,2	1.452	33,3	18.329	420,8	0,77	26	322
Emilia Romagna	27.587	618,6	4.068	91,2	711	15,9	10.695	239,9	0,86	14	205
Veneto	19.097	389,3	1.878	38,3	356	7,3	3.518	71,7	0,67	5	48
Toscana	10.067	269,9	1.015	27,2	280	7,5	2.871	77,0	0,76	6	58
Liguria	9.497	612,5	1.425	91,9	665	42,9	10.174	656,0	0,86	37	564
Lazio	7.643	130,0	688	11,7	453	7,7	4.241	72,1	0,65	5	47
Marche	6.716	440,3	995	65,2	173	11,3	2.543	166,7	0,62	7	104
Campania	4.755	82,0	405	7,0	153	2,6	1.312	22,6	0,75	2	17
Puglia	4.467	110,9	491	12,2	140	3,5	1.595	39,6	1,40	5	55
Trento	4.405	410,8	458	42,7	108	10,1	1.113	206,8	1,32	13	272
Sicilia	3.427	68,5	270	5,4	88	1,8	718	14,4	0,59	1	9
Friuli Venezia Giulia	3.240	266,6	329	27,1	102	8,4	1.042	85,8	0,83	7	71
Abruzzo	3.227	246,0	400	30,5	120	9,1	1.493	113,8	0,50	5	57
Bolzano	2.593	2.413,5	291	270,9	21	19,5	245	46,9	1,01	20	47
Umbria	1.430	162,1	75	8,5	18	2,0	NA	NA	0,83	2	NA
Sardegna	1.354	82,6	129	7,9	11	0,7	115	7,0	0,20	0	1
Valle dAosta	1.179	938,6	143	113,8	21	16,7	254	201,8	0,56	9	112
Calabria	1.157	59,4	96	4,9	23	1,2	NA	NA	0,38	0	NA
Molise	432	141,4	22	7,2	49	16,0	NA	NA	0,52	8	NA
Basilicata	399	70,9	27	4,8	13	2,3	NA	NA	1,74	4	NA
	Scale										
	Worst	Worst	Worst	Worst	Worst	Worst	Worst	Worst	2,0	100	1000
	Best	Best	Best	Best	Best	Best	Best	Best	0,0	0	0

(1) ρ_7 is the average of 7 consecutive ρ, but can still fluctuate. (2,3) EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

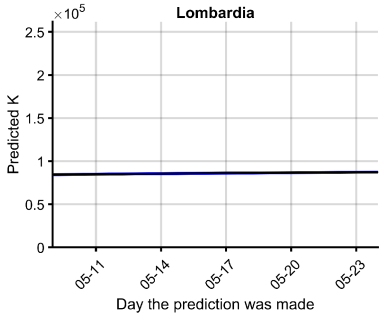
Disclaimer: estimated active cases and estimated 14-day attack rate are assessed by assuming a lethality of 1 % (see report from 20 to 24 April, #37-41). This value can change in countries where suspicious deaths are reported as well (real values would be lower) and in countries where incidence among elderly people was minor (real values would be higher).

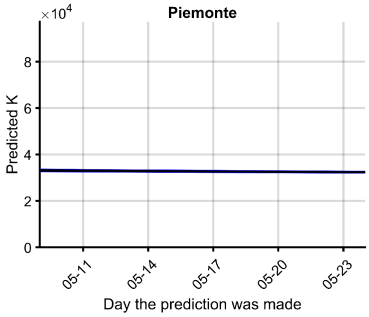
Long-term predictions

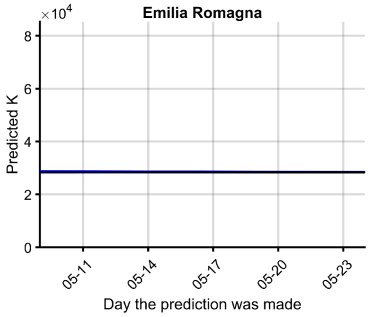


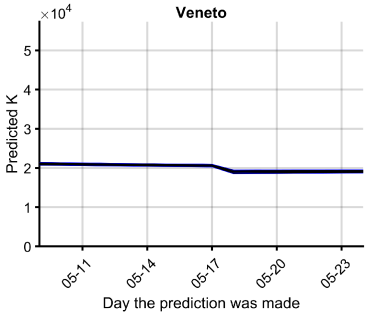
[1] **Spain:** Historical series have not been updated. Therefore, regional analysis is not shown

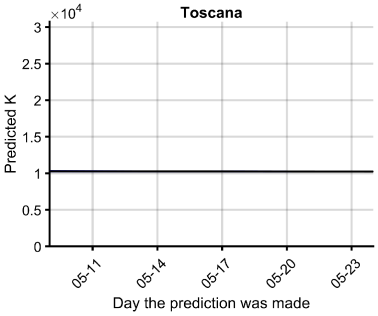
2020-05-25

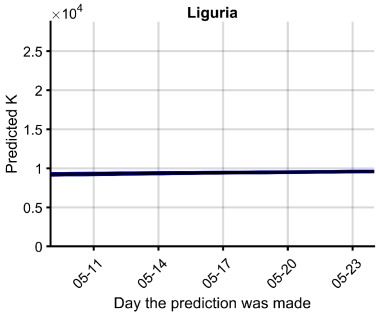

Lombardia
×10^5
Predicted K
Day the prediction was made

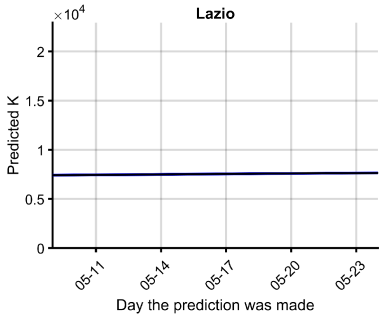

Piemonte
×10^4
Predicted K
Day the prediction was made

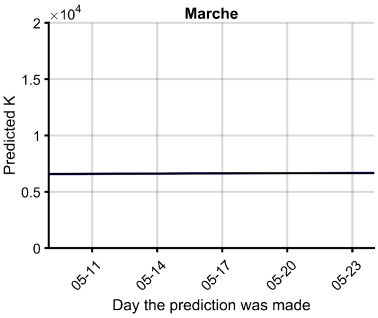

Emilia Romagna
×10^4
Predicted K
Day the prediction was made

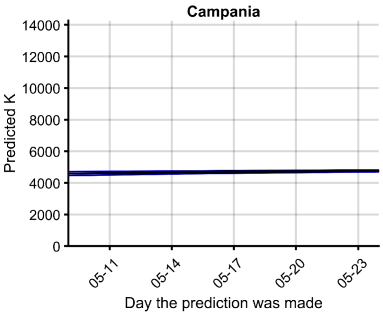

Veneto
×10^4
Predicted K
Day the prediction was made

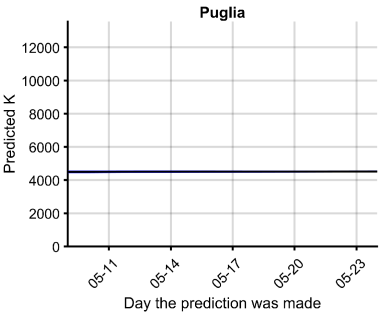

Toscana
×10^4
Predicted K
Day the prediction was made

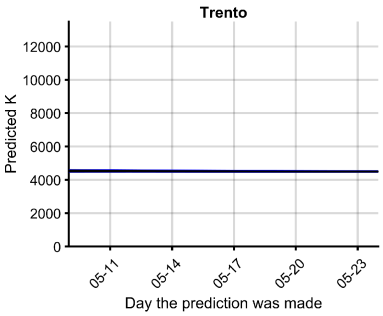

Liguria
×10^4
Predicted K
Day the prediction was made

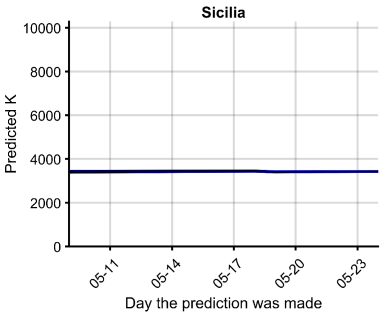

Lazio
×10^4
Predicted K
Day the prediction was made

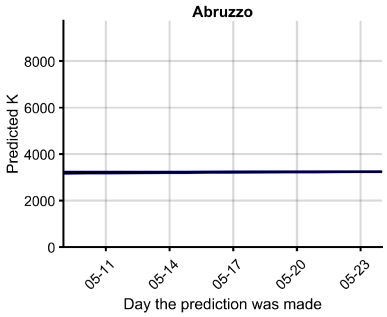

Marche
×10^4
Predicted K
Day the prediction was made

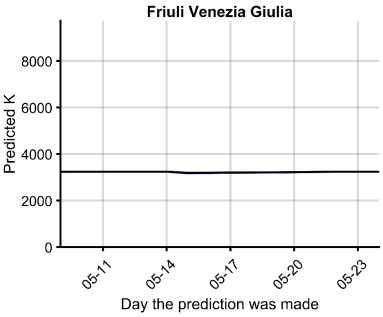

Campania
Predicted K
Day the prediction was made

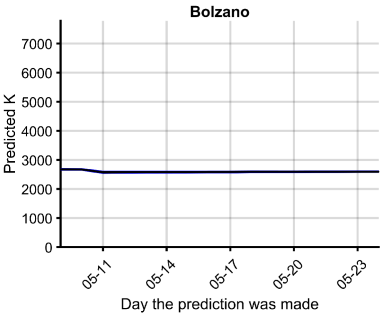

Puglia
Predicted K
Day the prediction was made

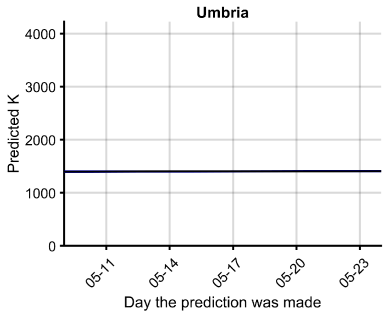

Trento
Predicted K
Day the prediction was made

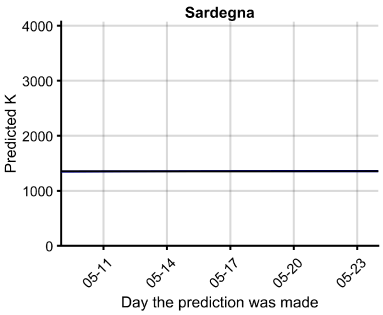

Sicilia
Predicted K
Day the prediction was made

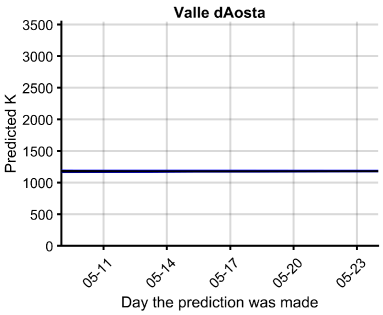

Abruzzo
Predicted K
Day the prediction was made

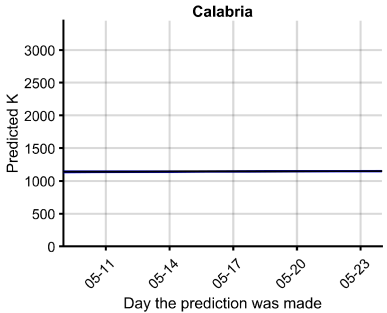

Friuli Venezia Giulia
Predicted K
Day the prediction was made

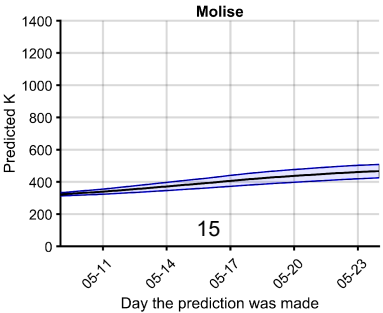

Bolzano
Predicted K
Day the prediction was made

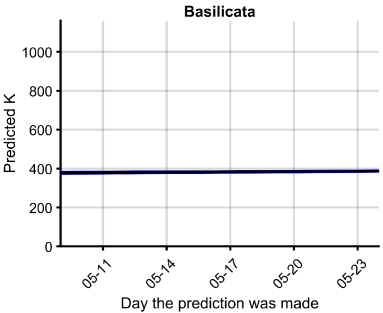

Umbria
Predicted K
Day the prediction was made


Sardegna
Predicted K
Day the prediction was made


Valle dAosta
Predicted K
Day the prediction was made


Calabria
Predicted K
Day the prediction was made

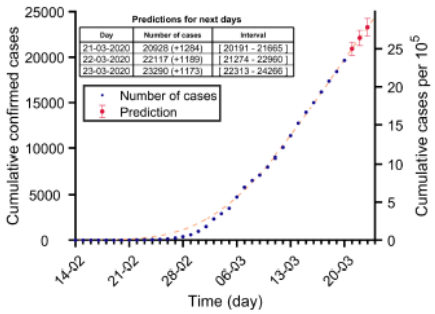

Molise
Predicted K
Day the prediction was made

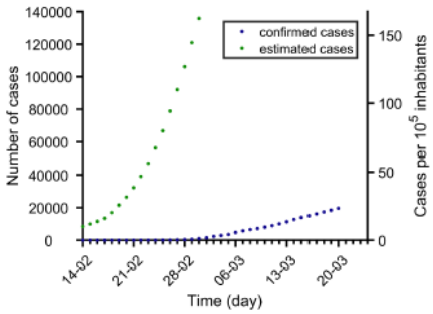

Basilicata
Predicted K
Day the prediction was made

Legend: Countries' reports details

Iran 20-03-2020. Population: 83.7M. Current cumulated incidence: $23/10^5$

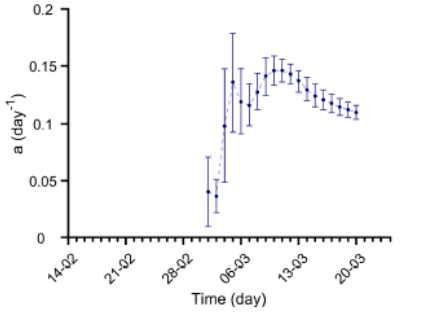
Confirmed cases: data (blue), model fitted (dashed line), predictions (red points and table)

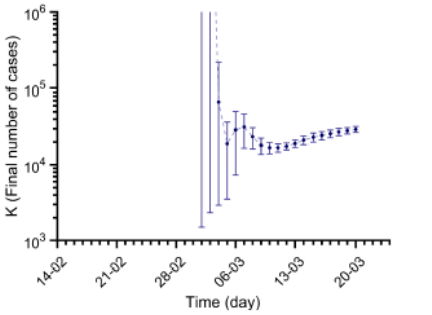




Estimated cases using death rate (see Methods)

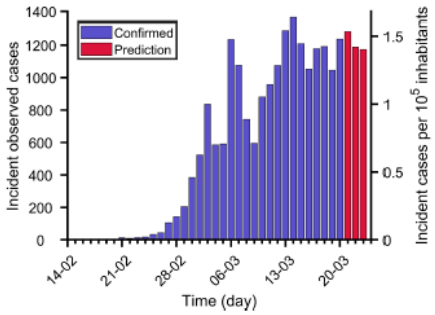
Fitted *a* value using points prior to each date

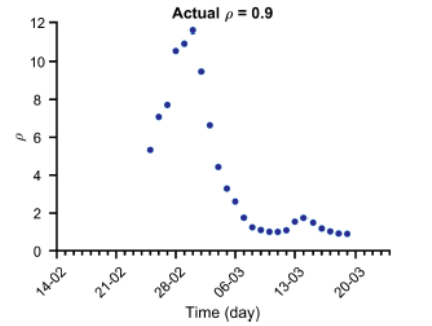




Fitted *K* value using points prior to each date

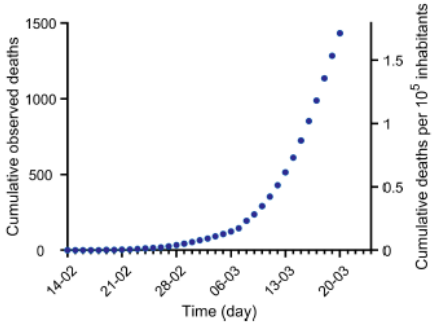
Reported and predicted new cases

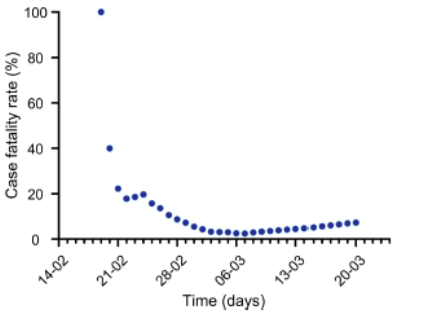




Evolution of ρ, a parameter related with Reproduction number (see Methods)

Reported deaths



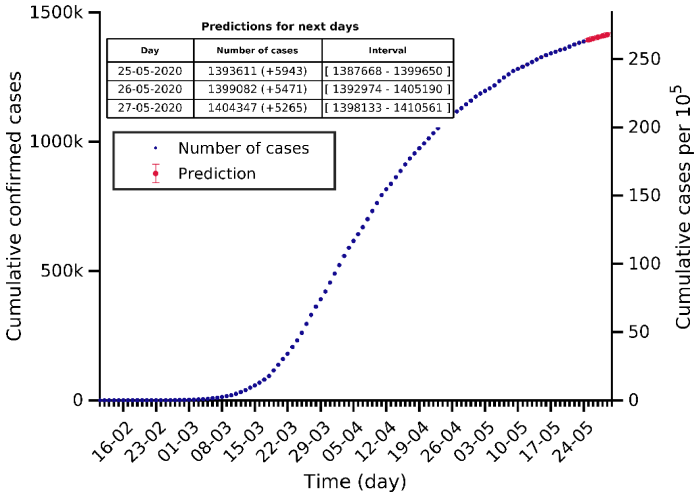


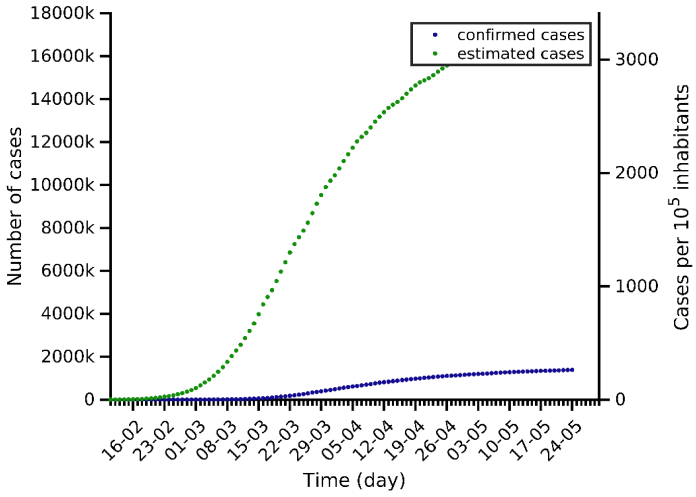
Deaths / cumulated reported cases

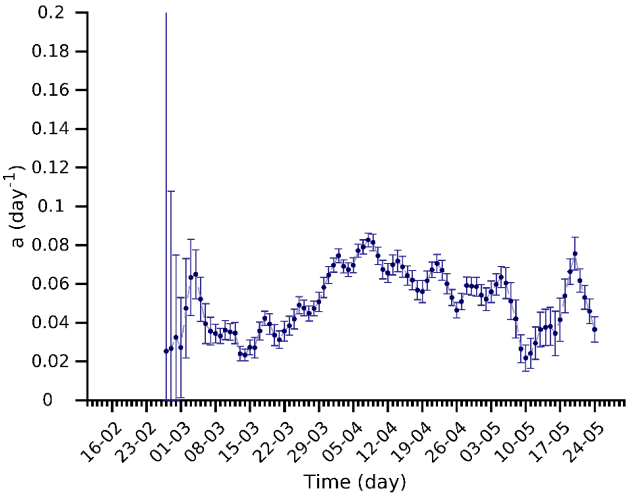
(1) Analysis and prediction of COVID-19 for EU+EFTA+UK

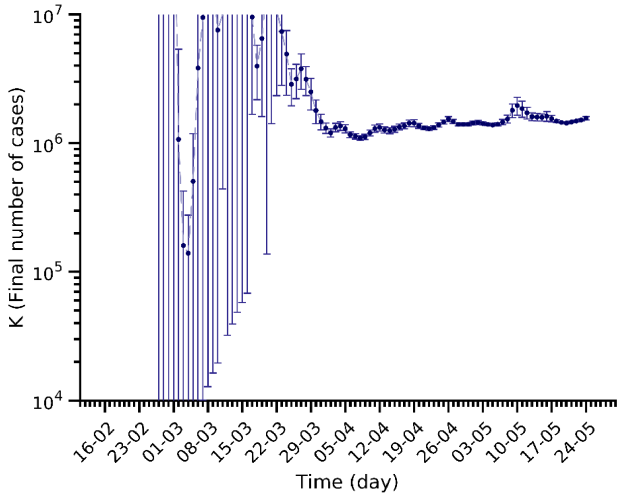
Data obtained from https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

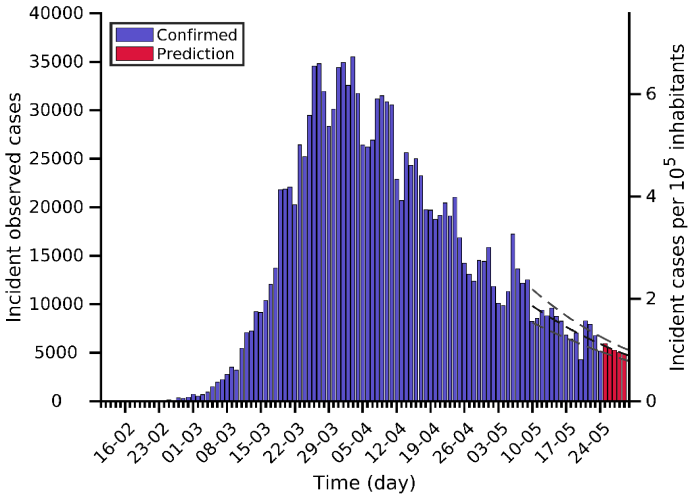
EU+EFTA+UK 24-05-2020. Population: 527.9M. Current cumulated incidence: 263/1

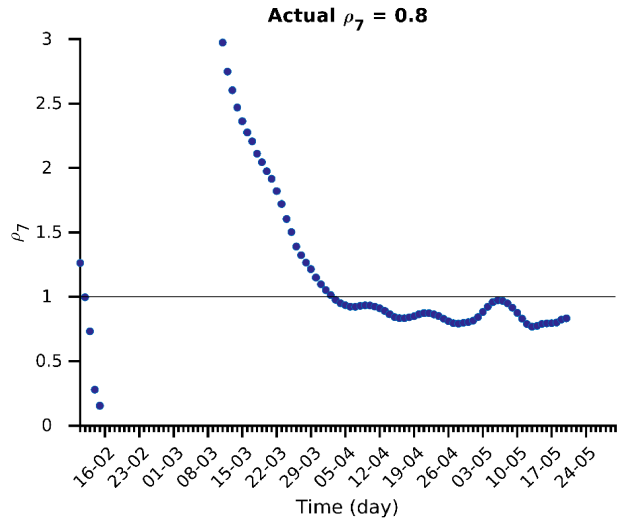


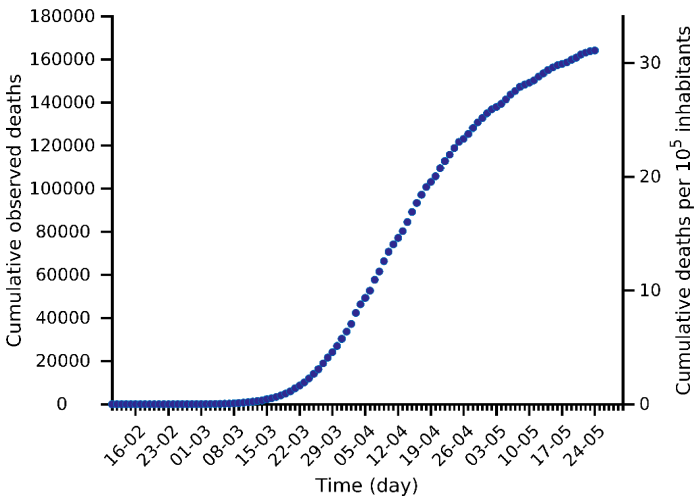


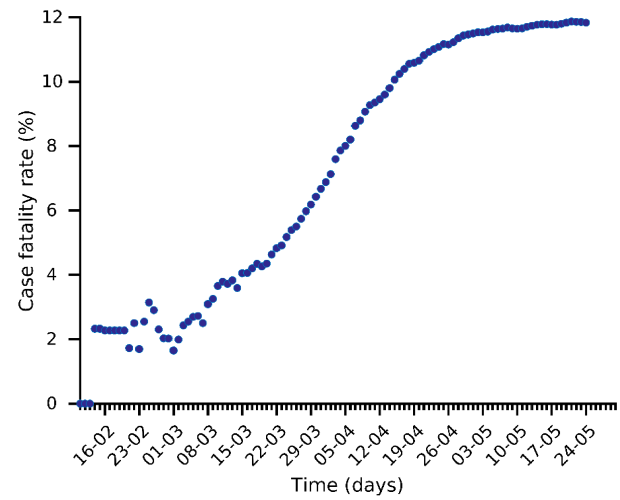




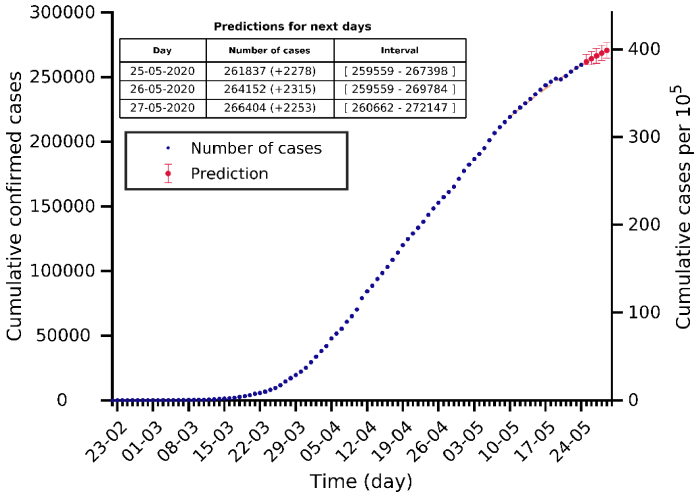


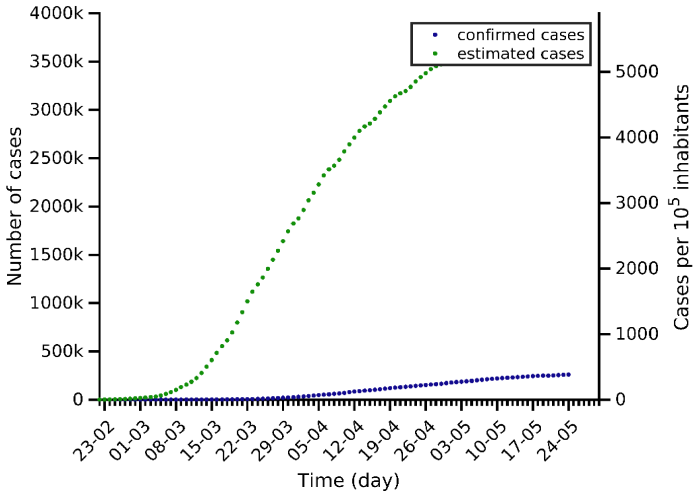


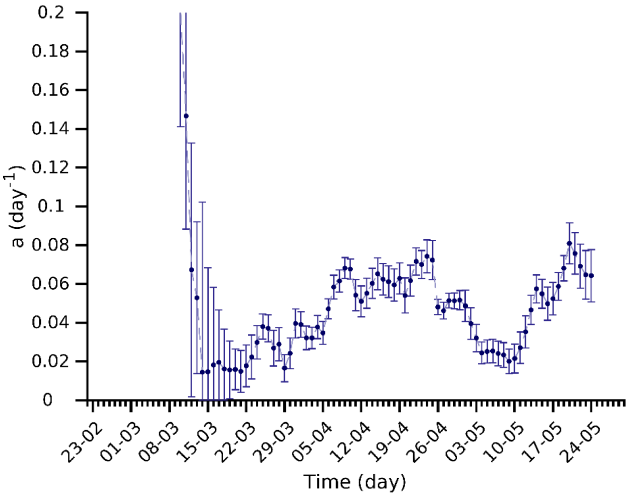


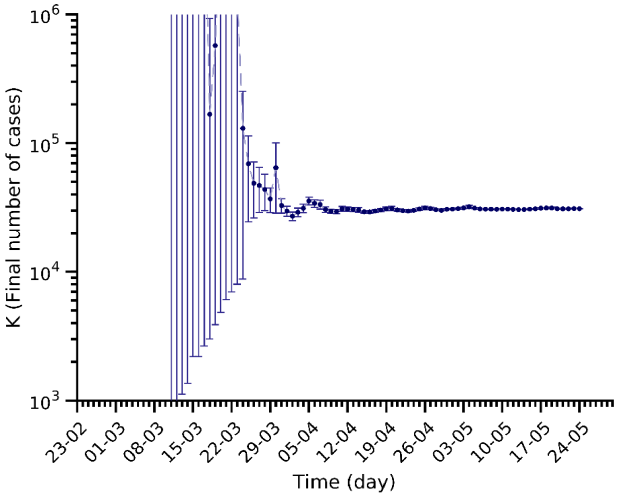


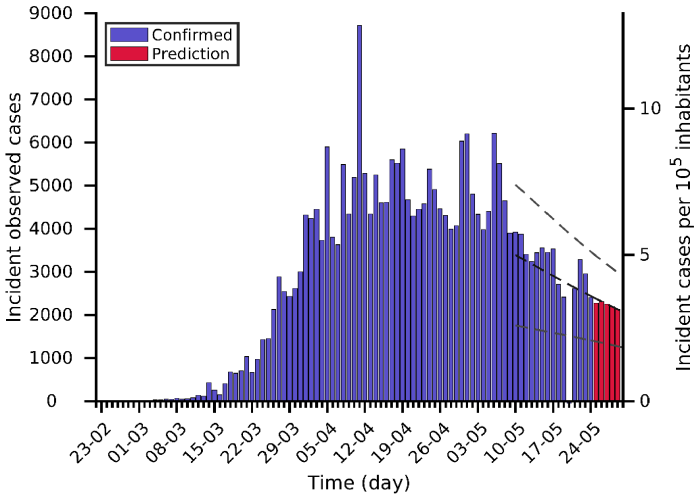
UK 24-05-2020. Population: 67.9M. Current cumulated incidence: $382/10^5$

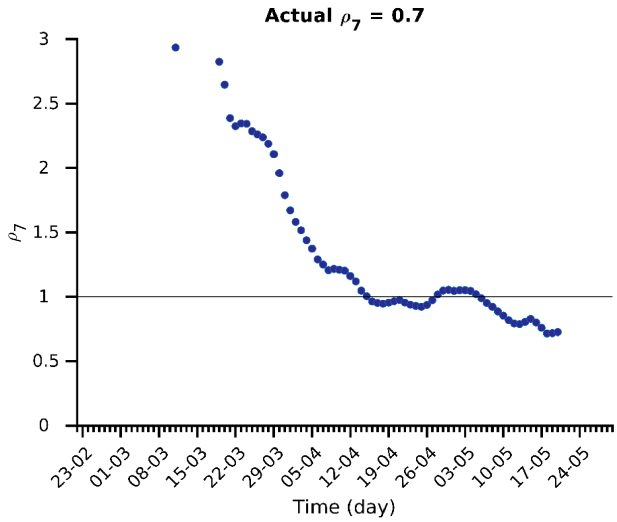


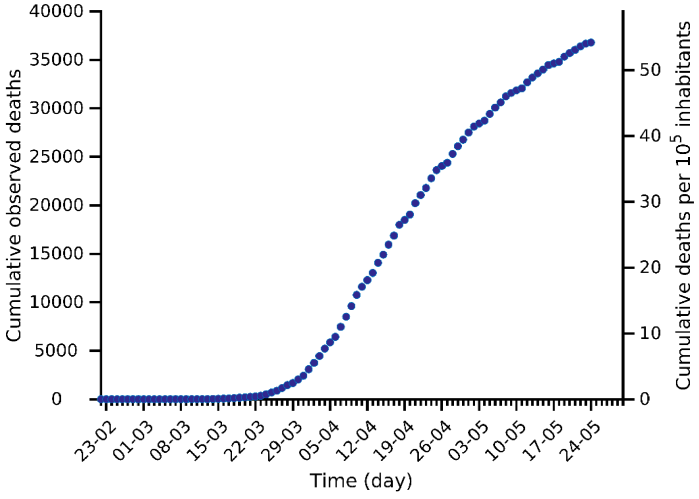


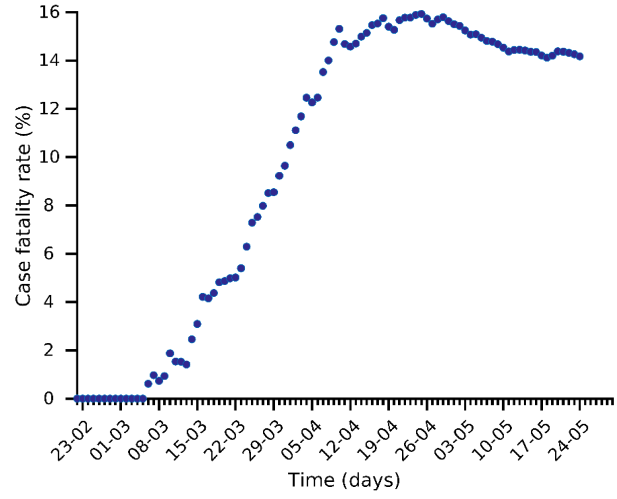




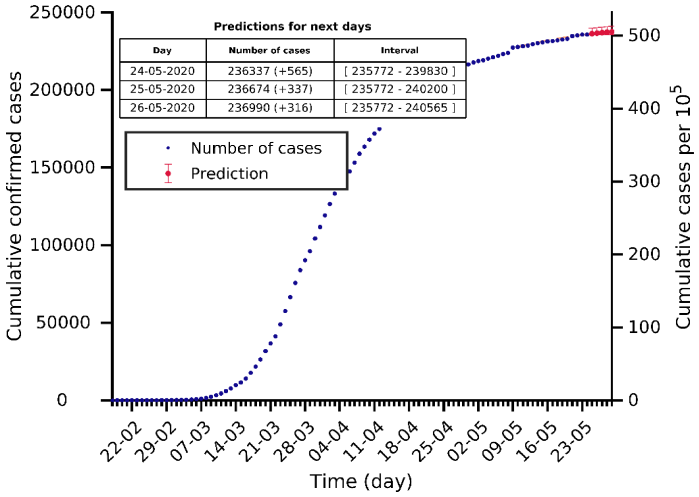


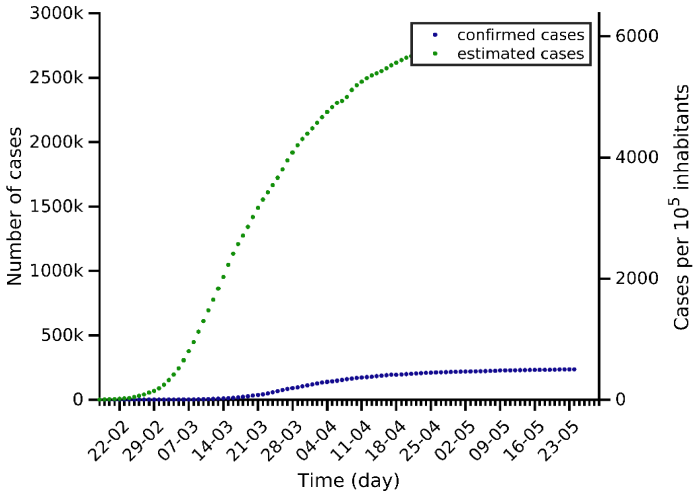


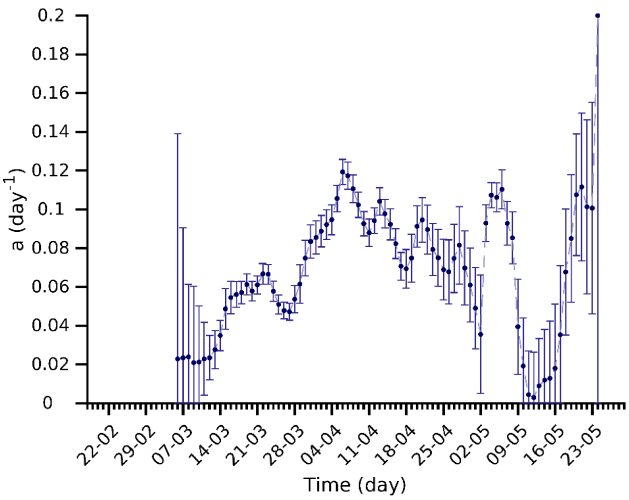


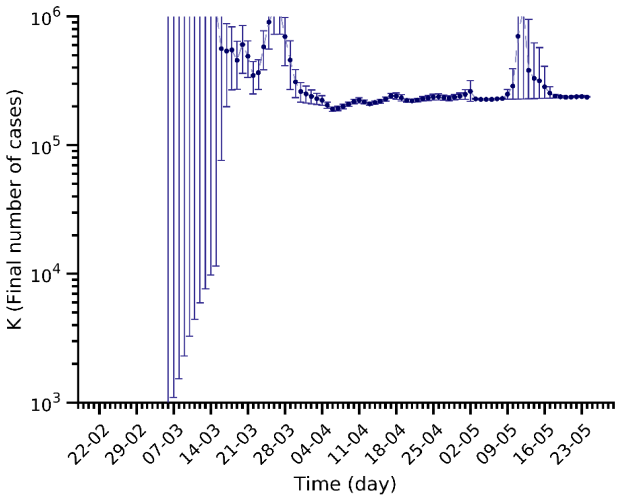


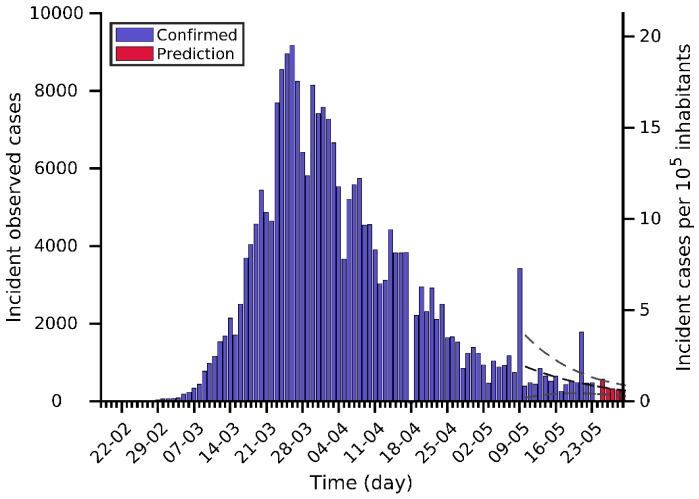
Spain 23-05-2020. Population: 47.0M. Current cumulated incidence: 501/10⁵

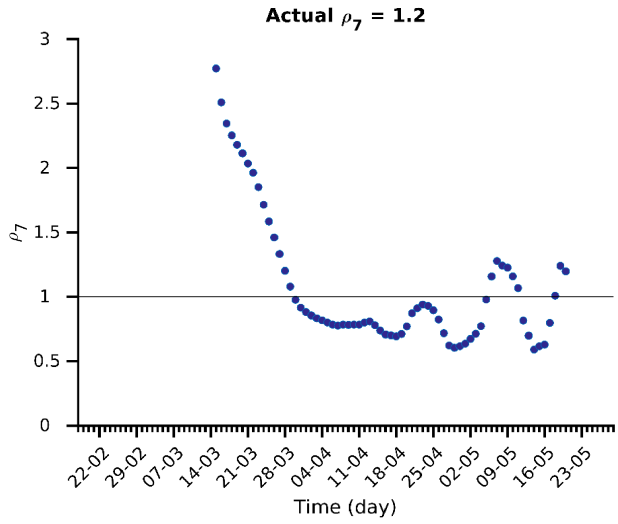


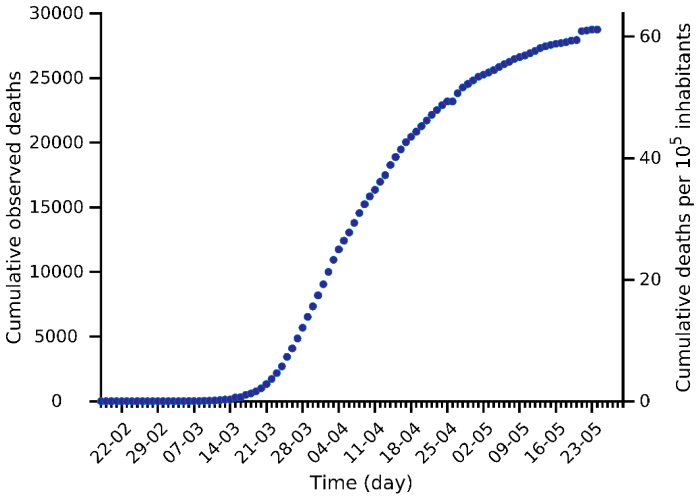


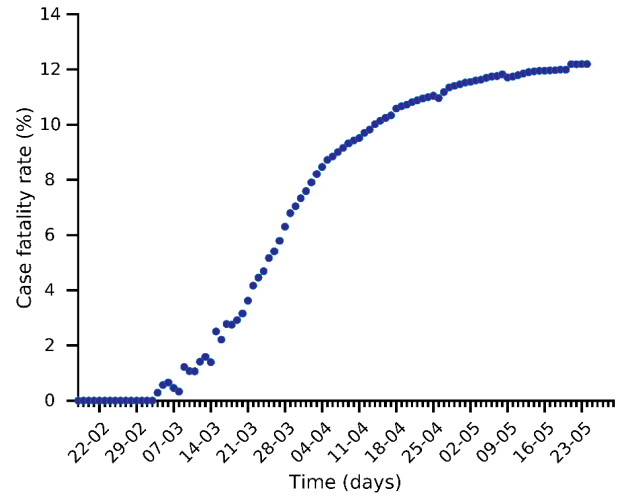




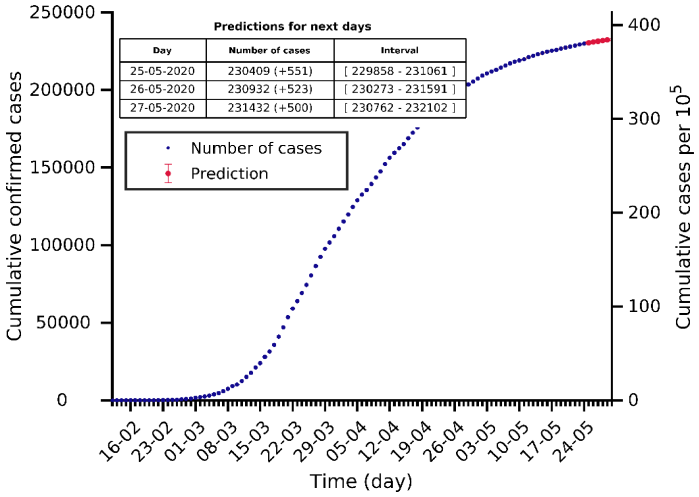


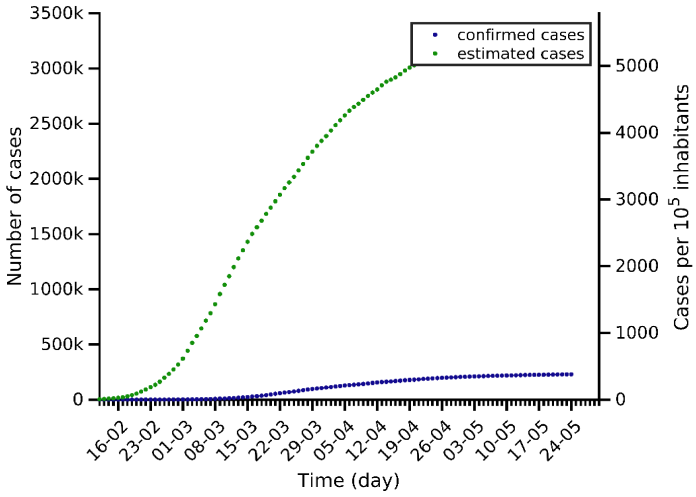


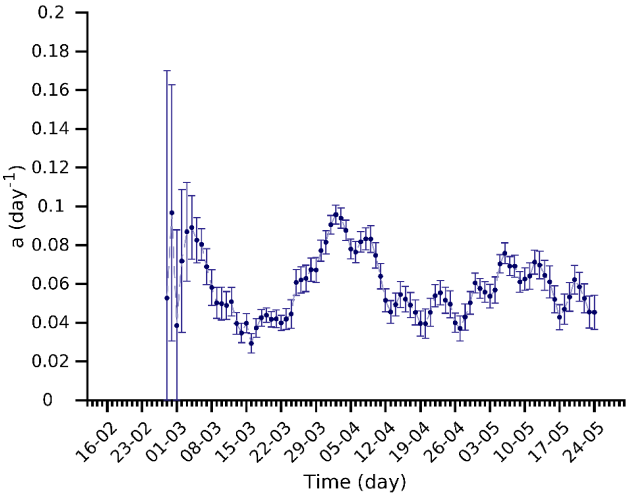


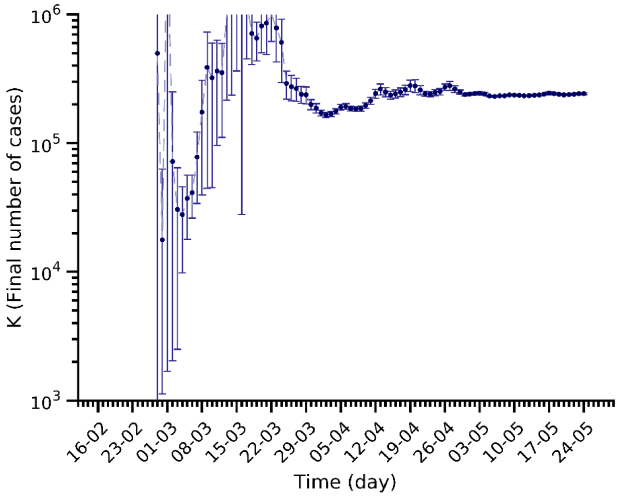


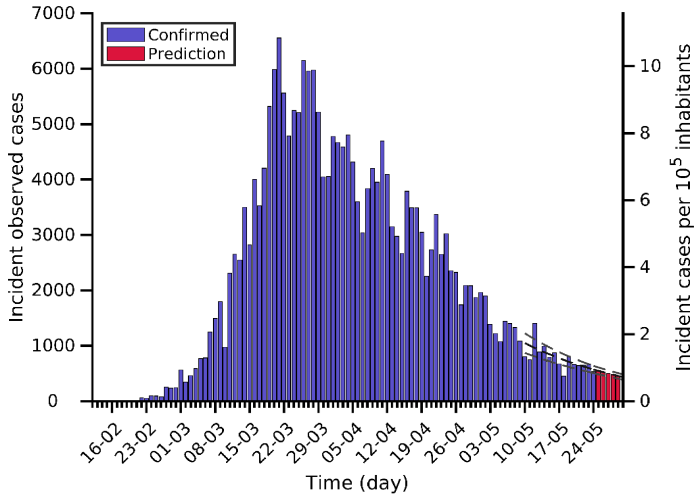
Italy 24-05-2020. Population: 60.5M. Current cumulated incidence: $380/10^5$

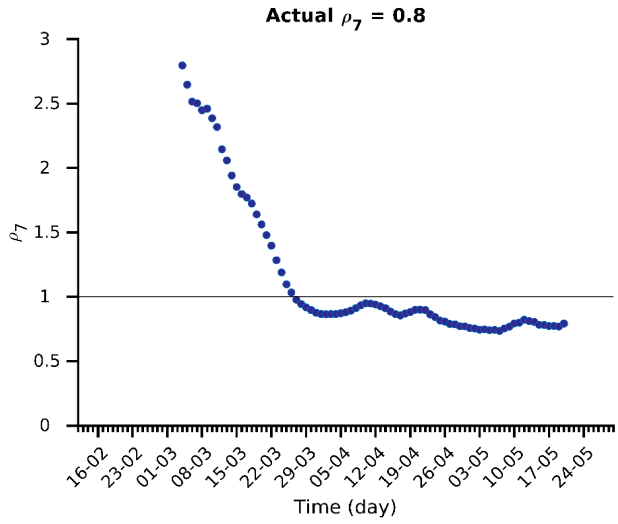


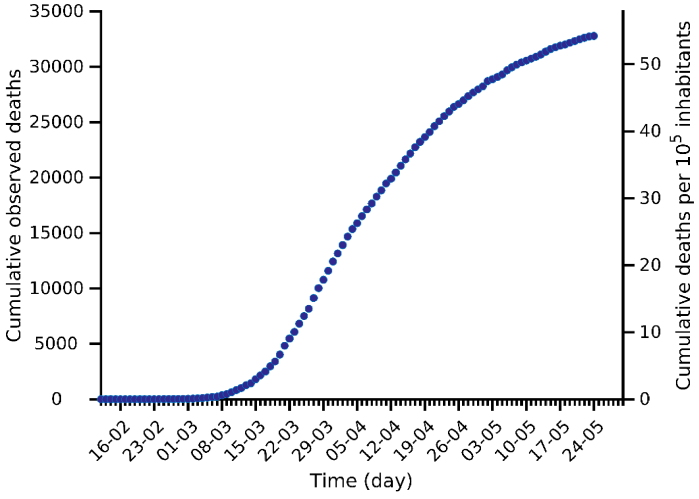


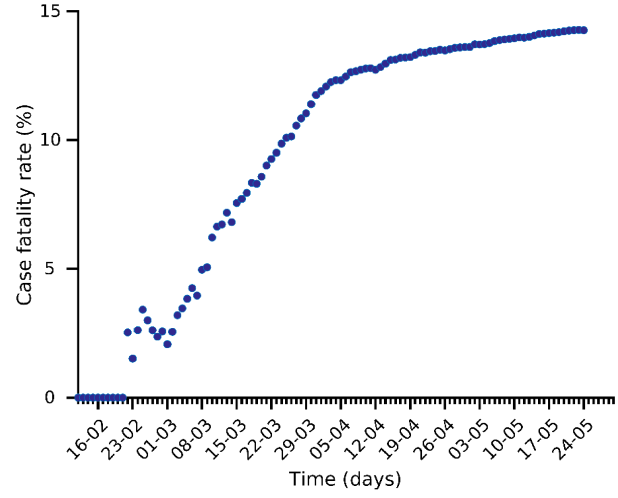




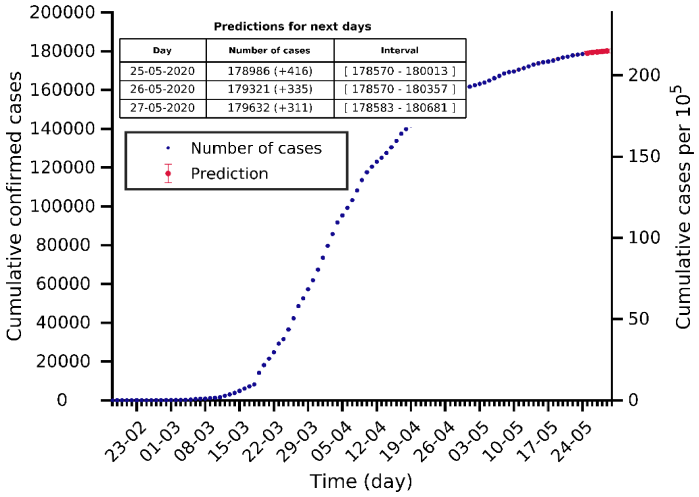


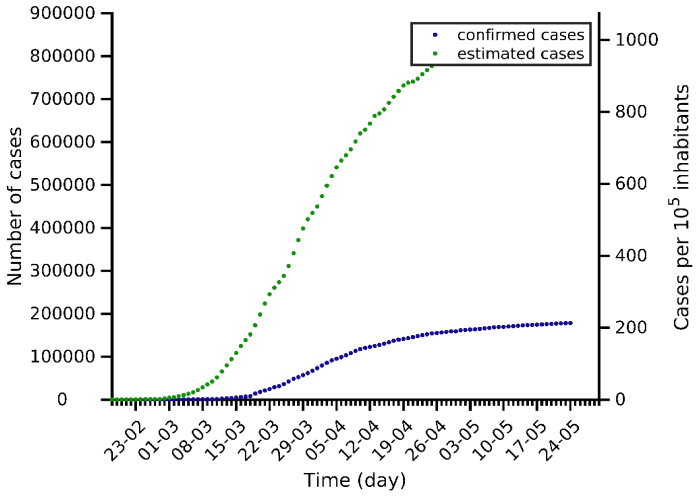


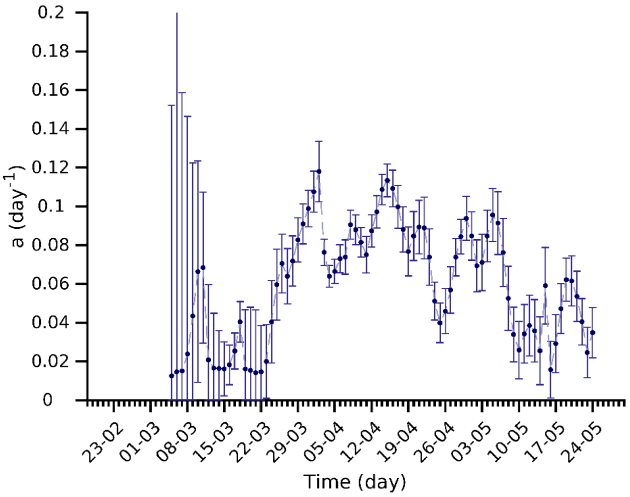


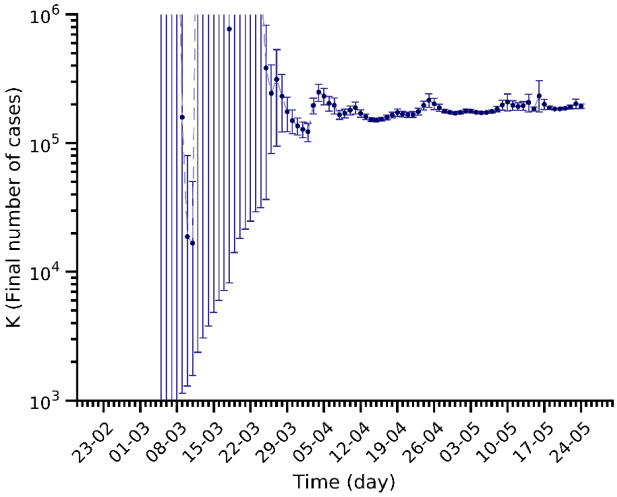


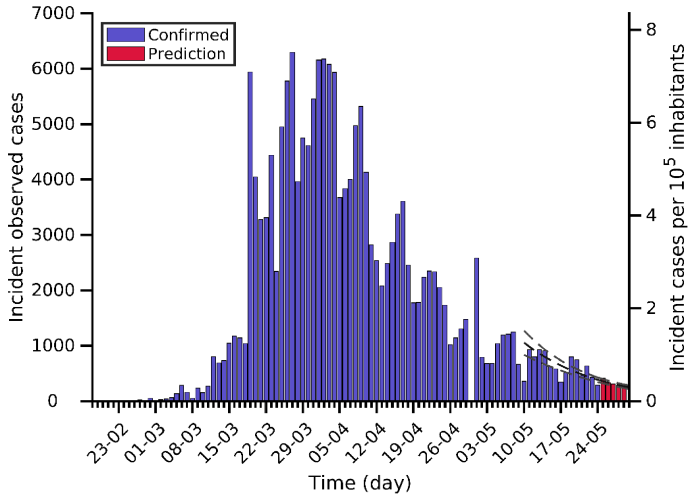
Germany 24-05-2020. Population: 83.8M. Current cumulated incidence: 213/10^5

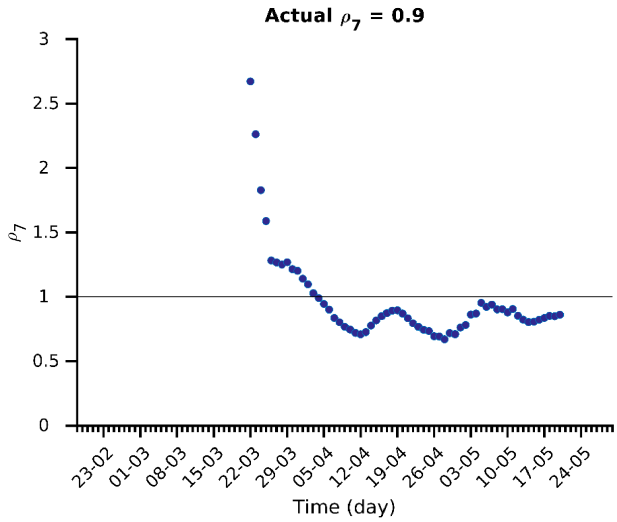


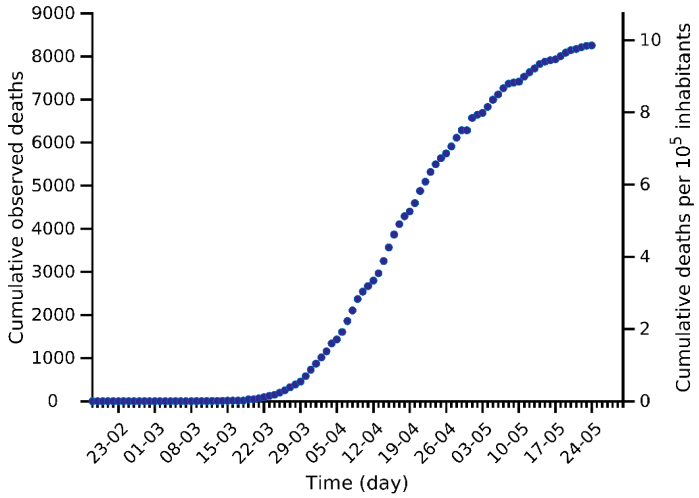


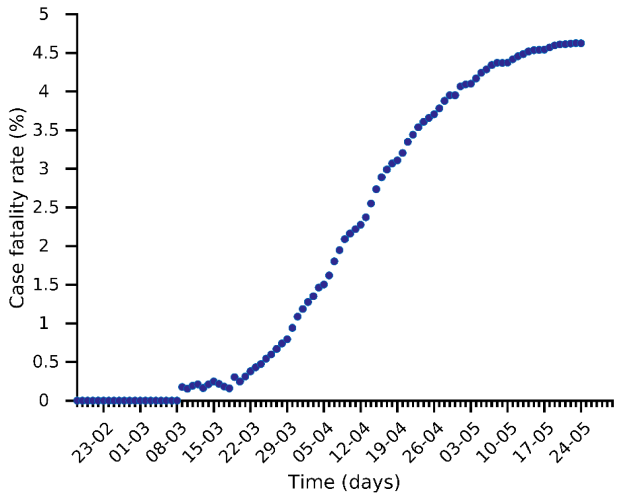




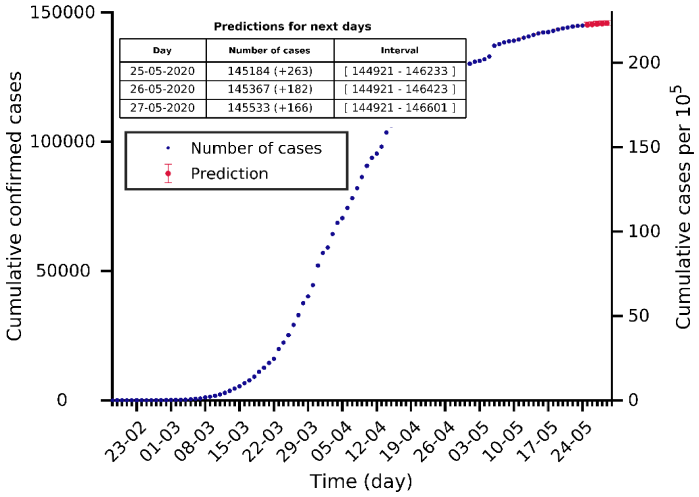


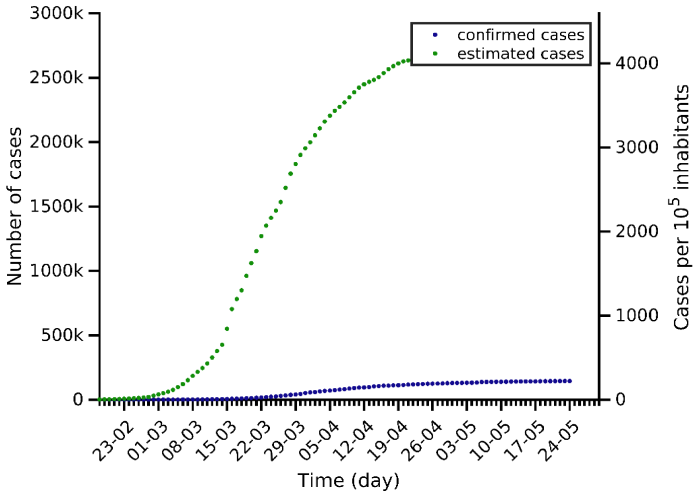


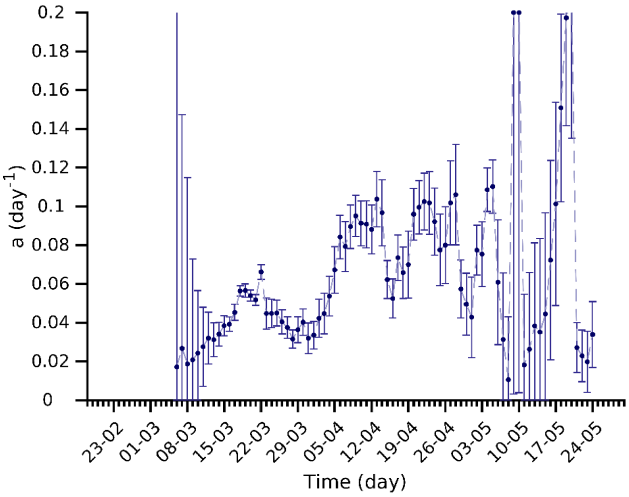


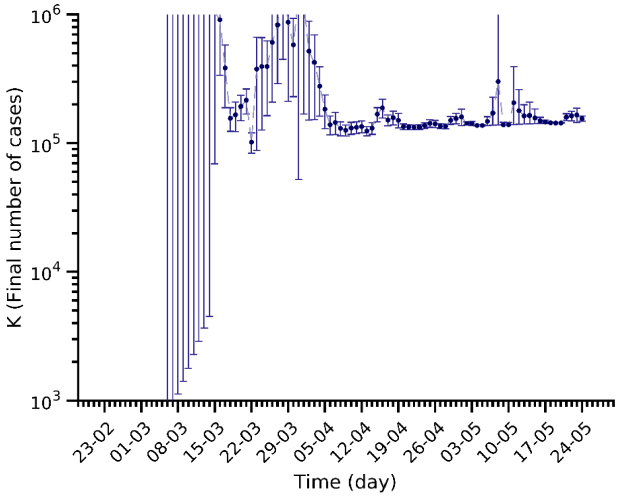


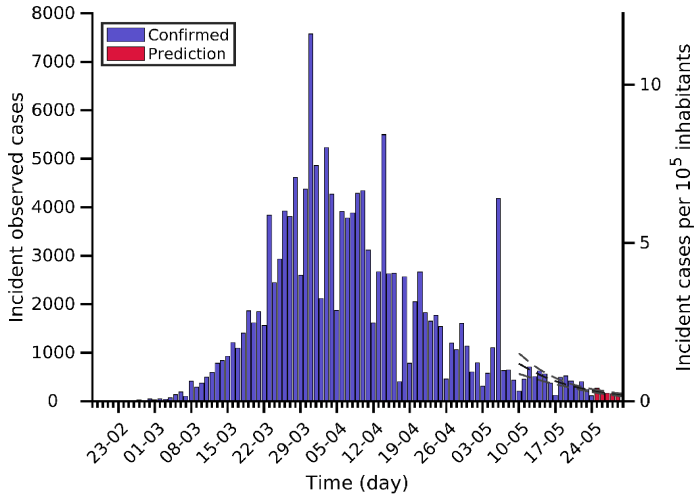
France 24-05-2020. Population: 65.3M. Current cumulated incidence: 222/10^5

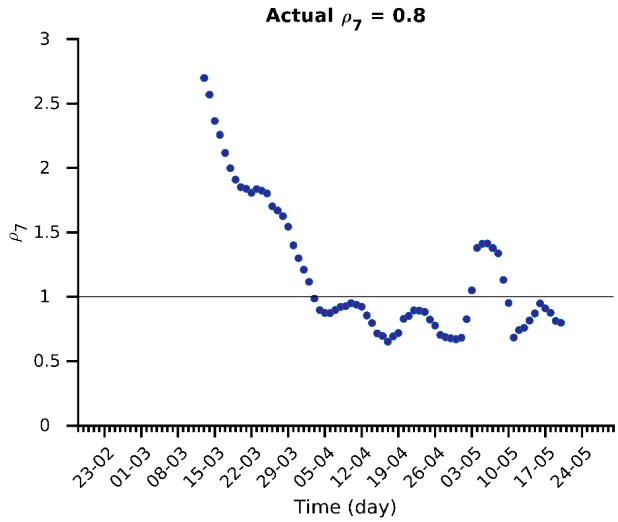


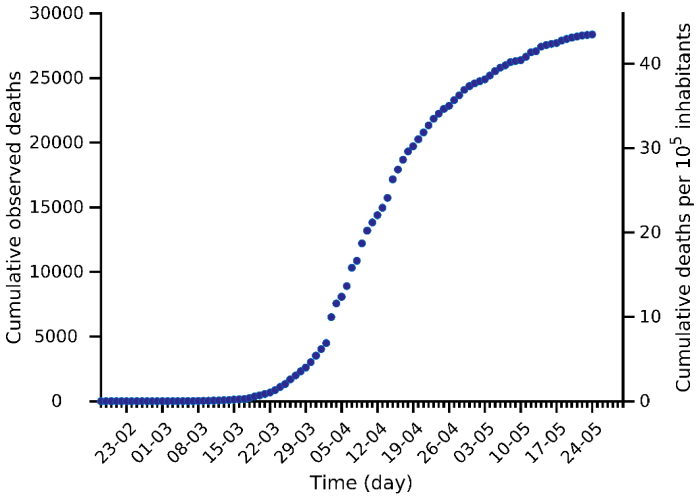


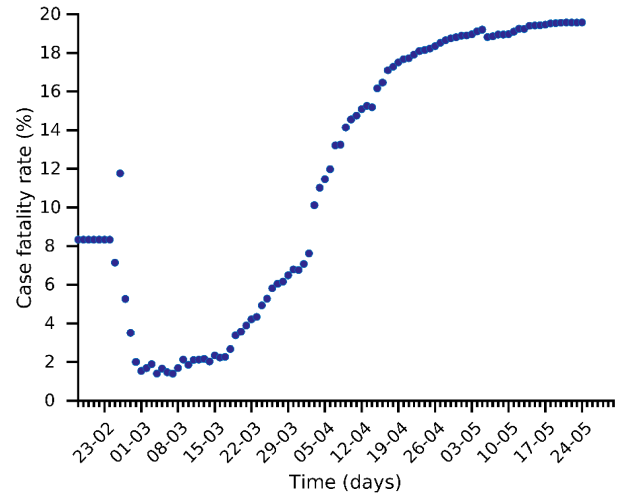




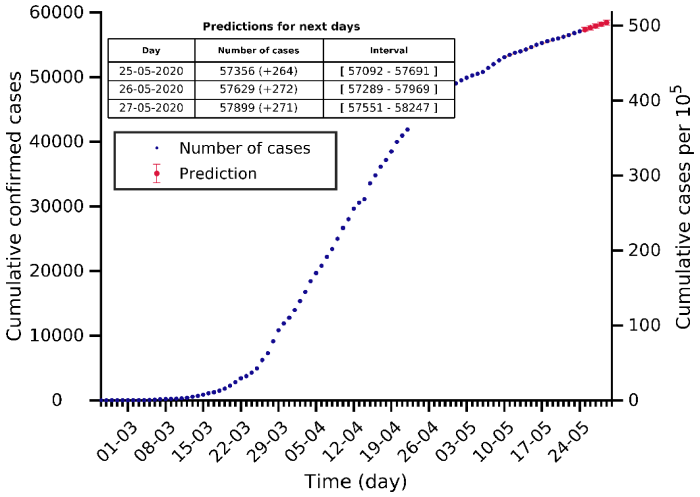


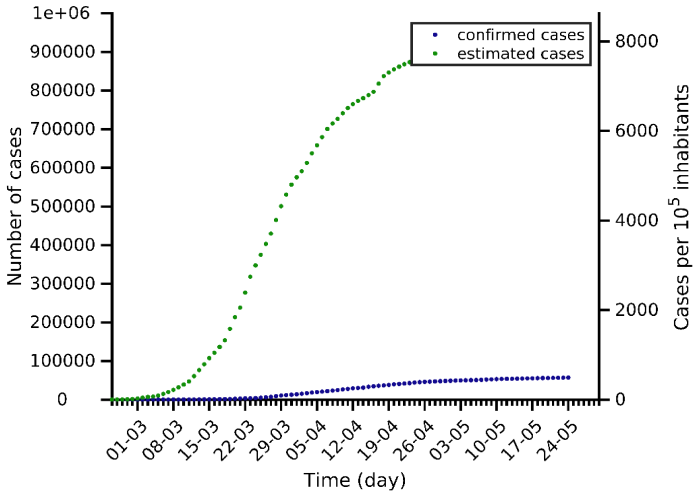


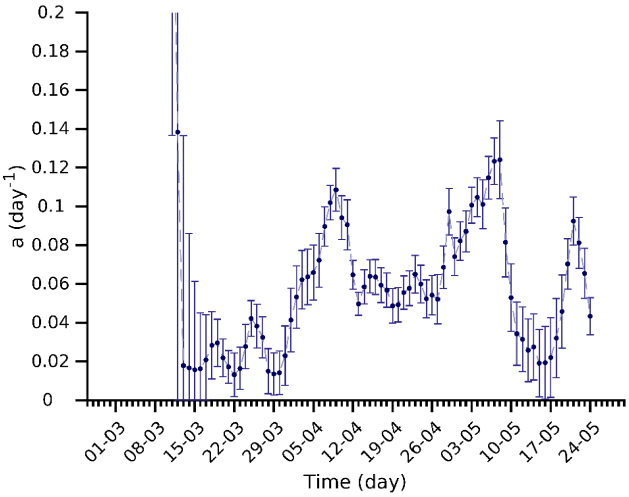


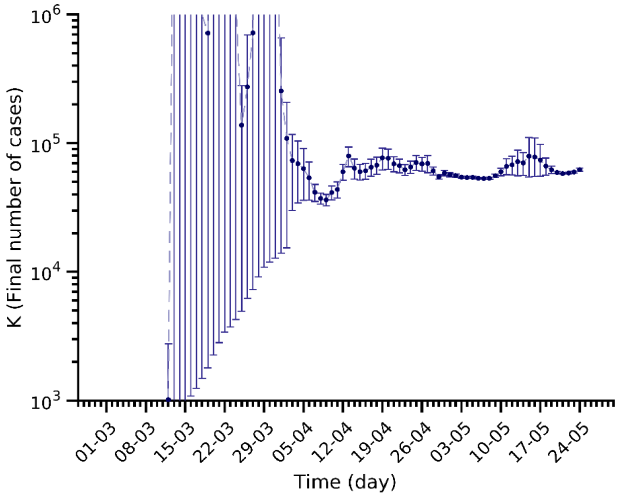


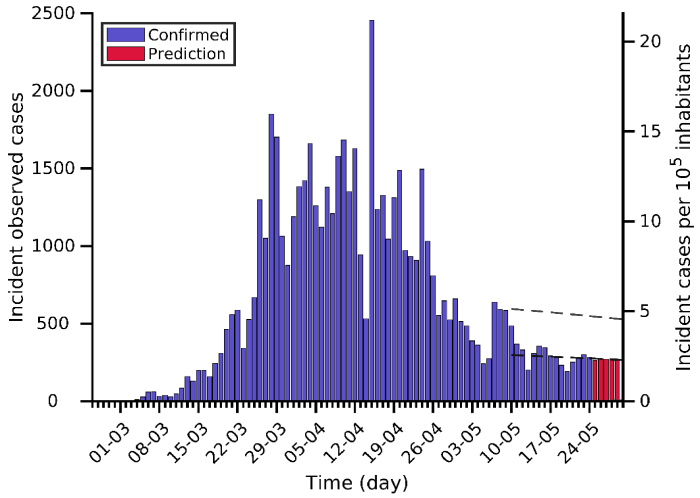
Belgium 24-05-2020. Population: 11.6M. Current cumulated incidence: 493/10^5

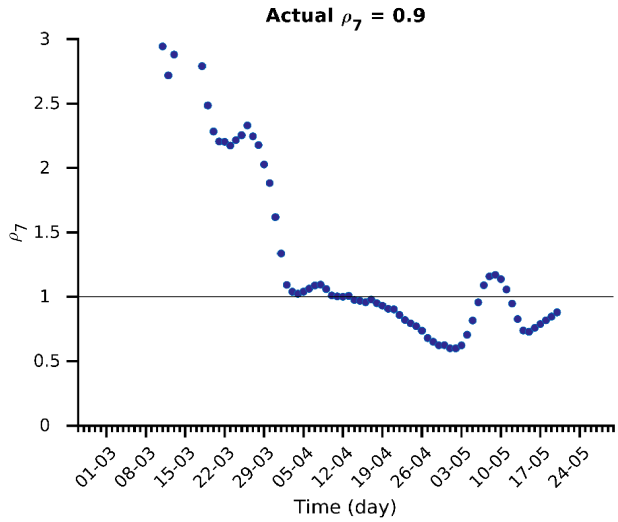


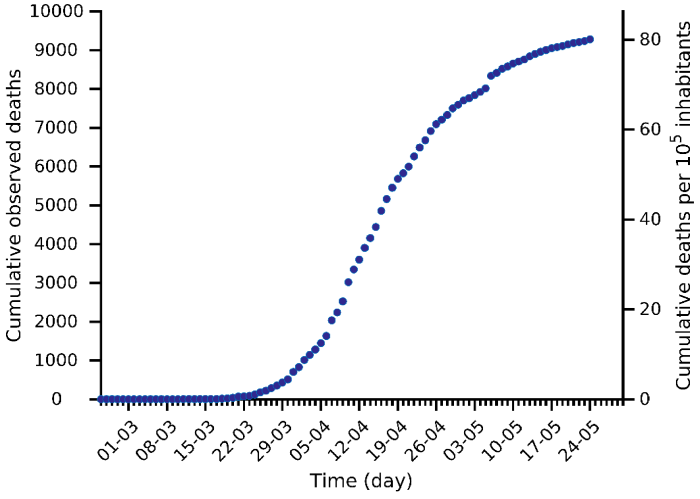


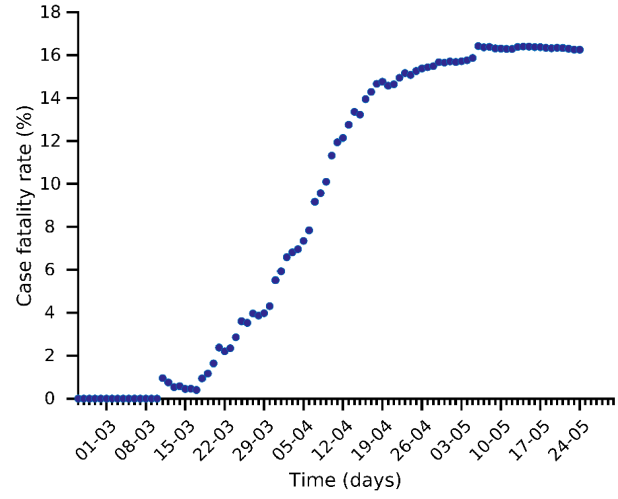




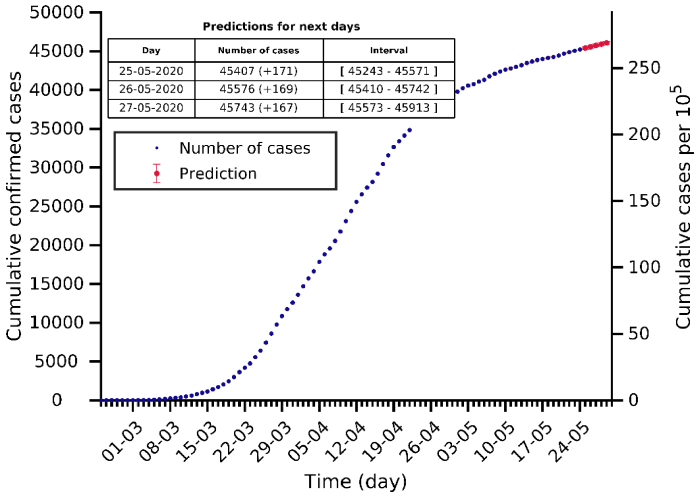


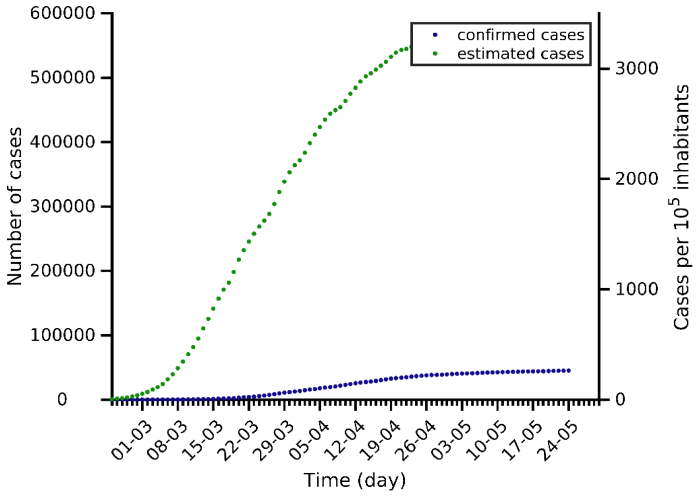


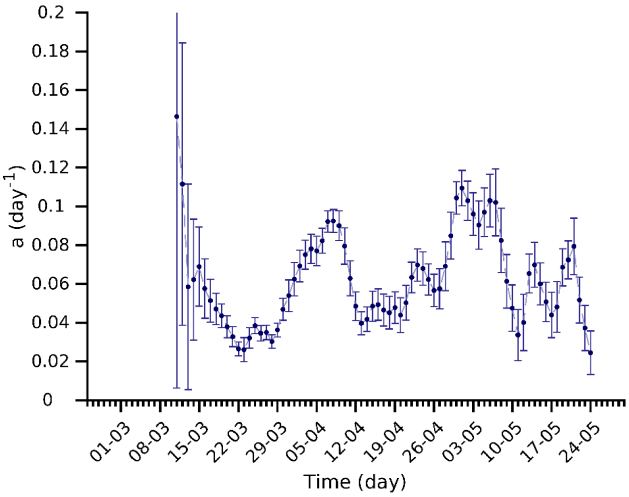


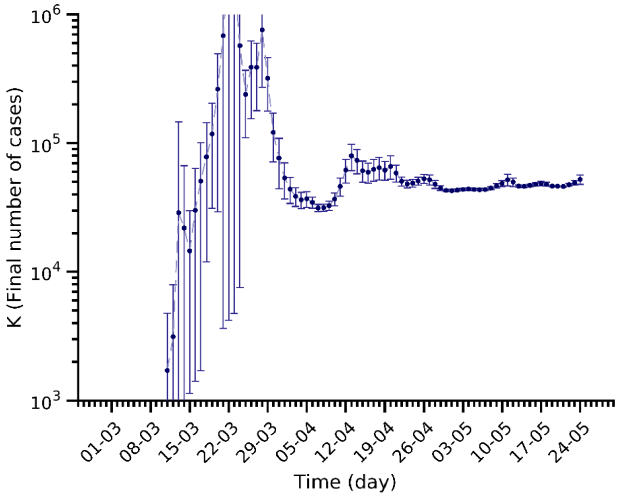


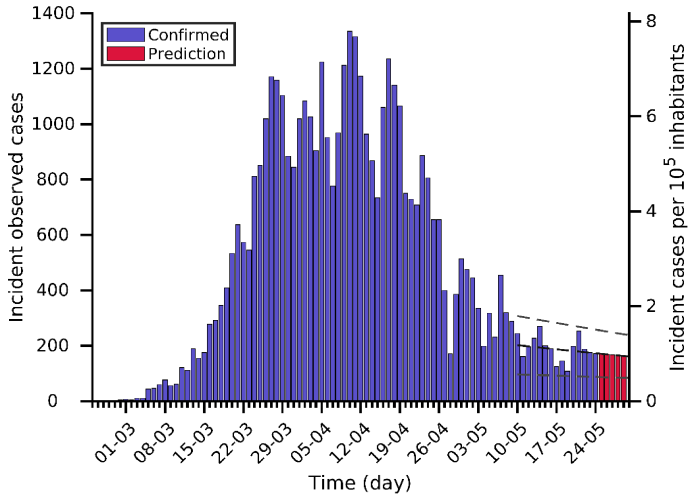
Netherlands 24-05-2020. Population: 17.1M. Current cumulated incidence: 264/10^5

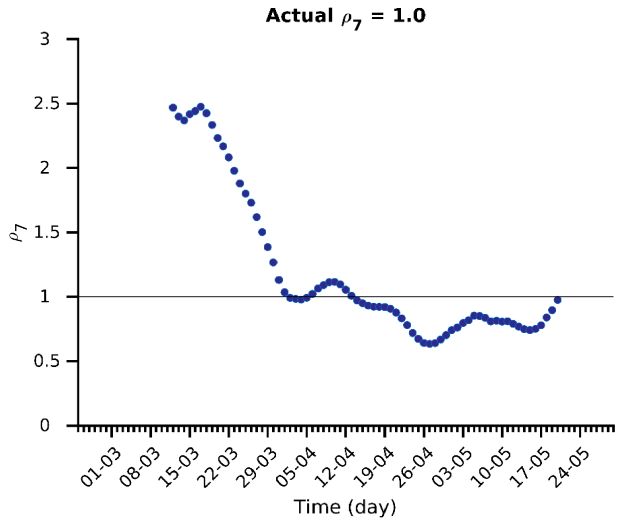


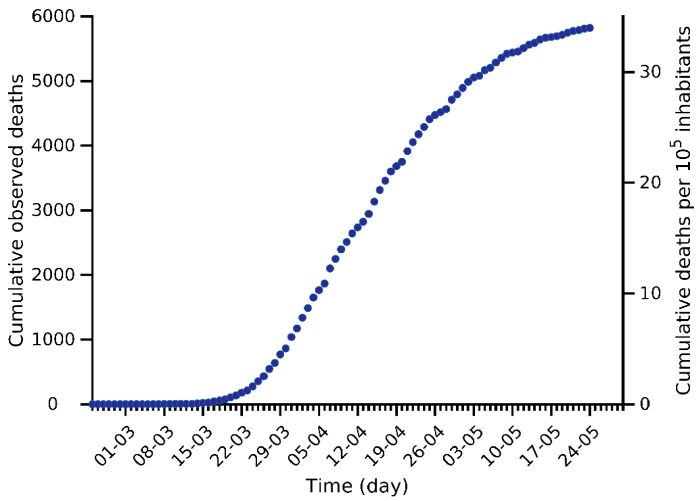


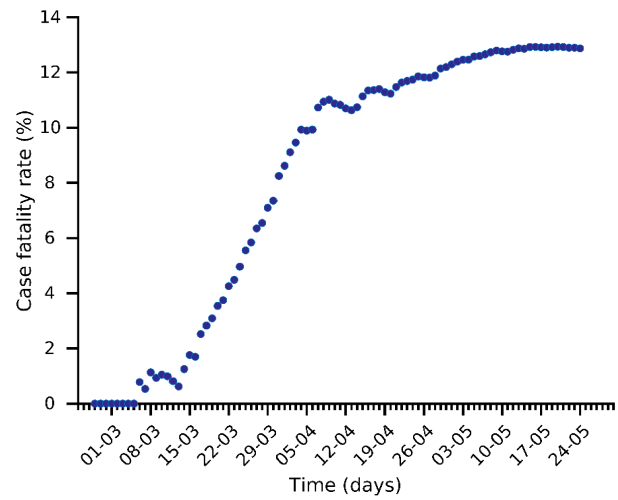




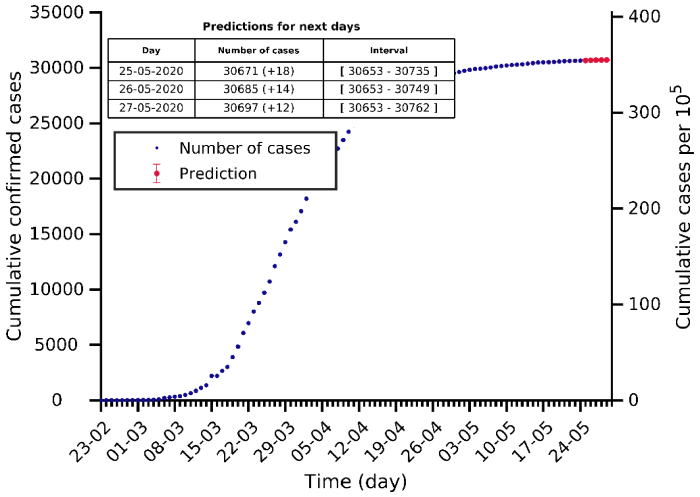


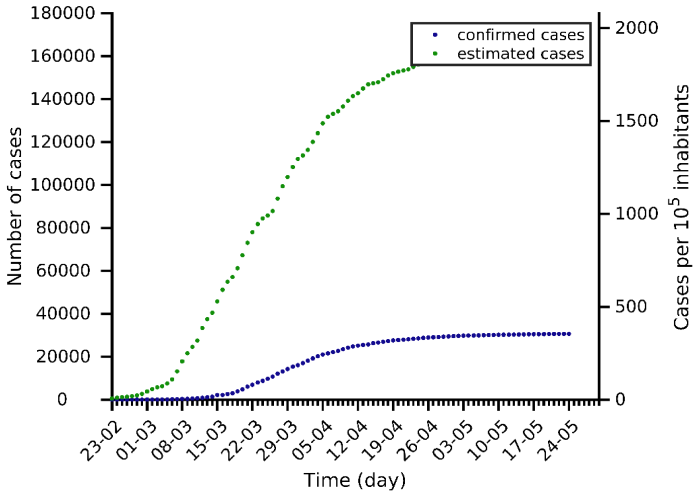


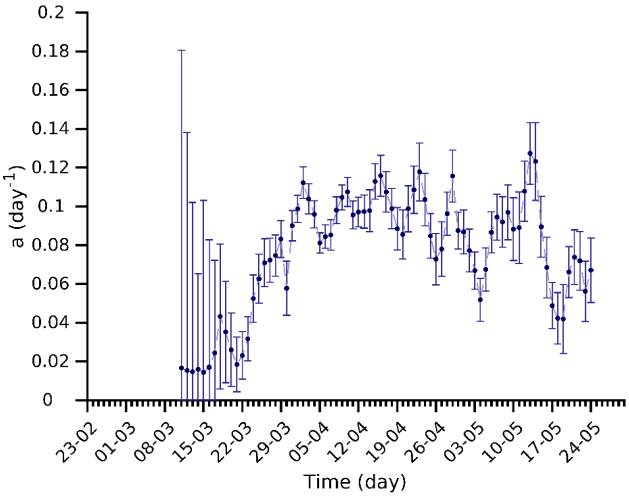


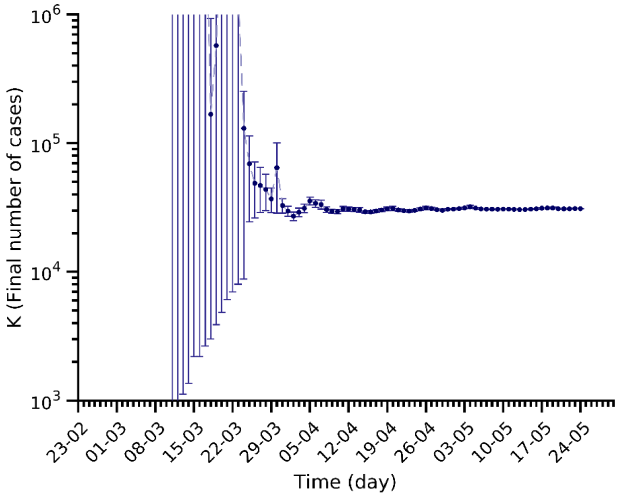


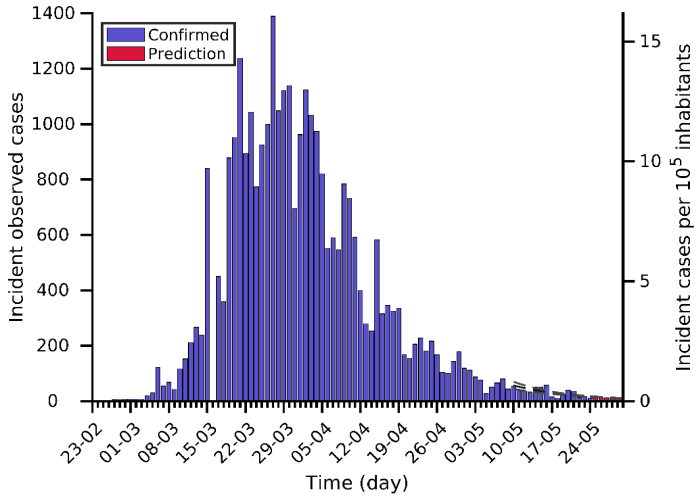
Switzerland 24-05-2020. Population: 8.7M. Current cumulated incidence: 354/10⁵

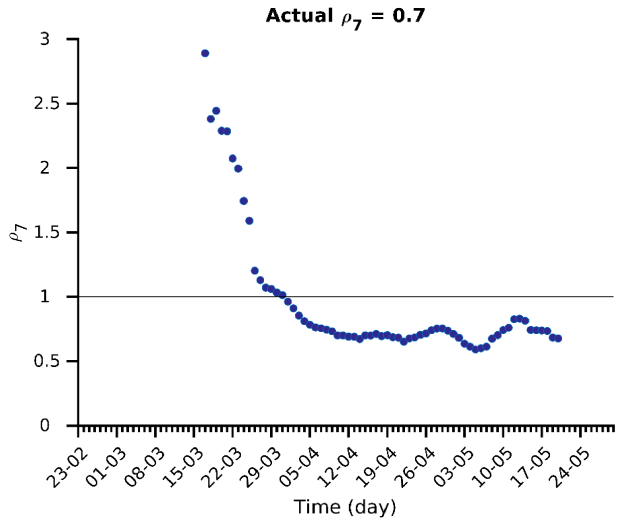


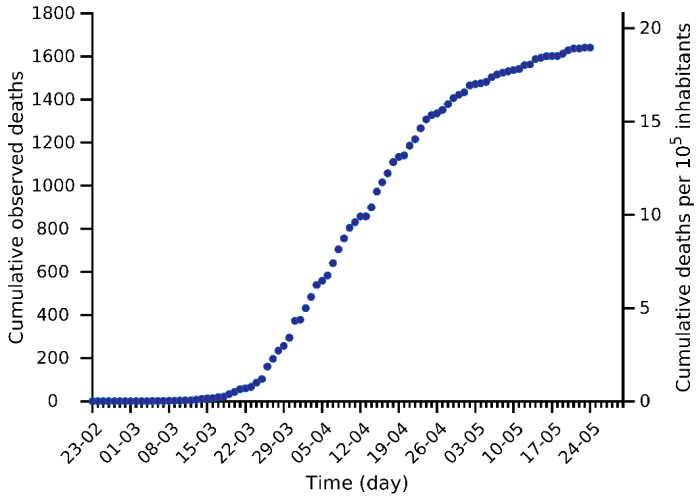


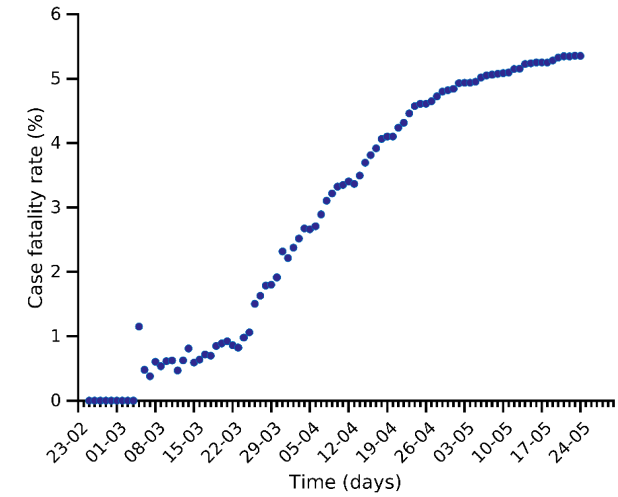




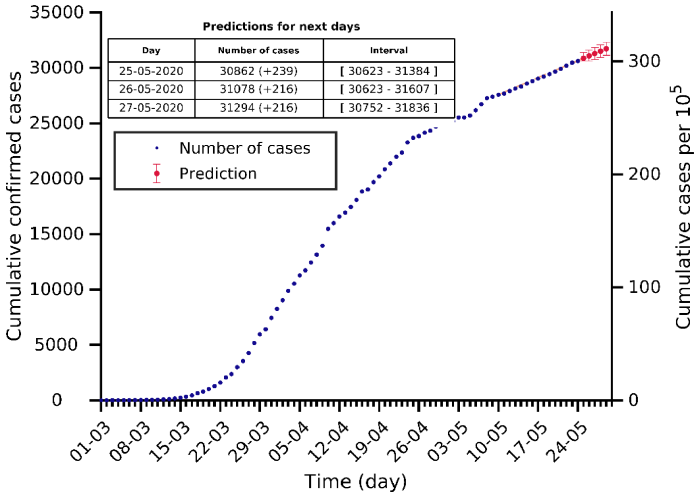


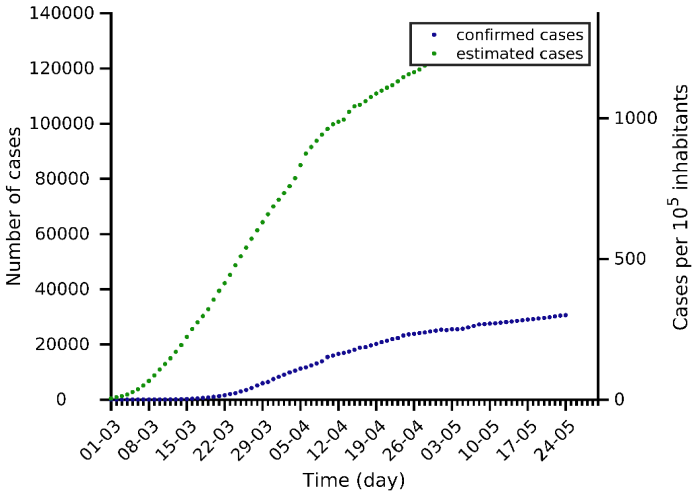


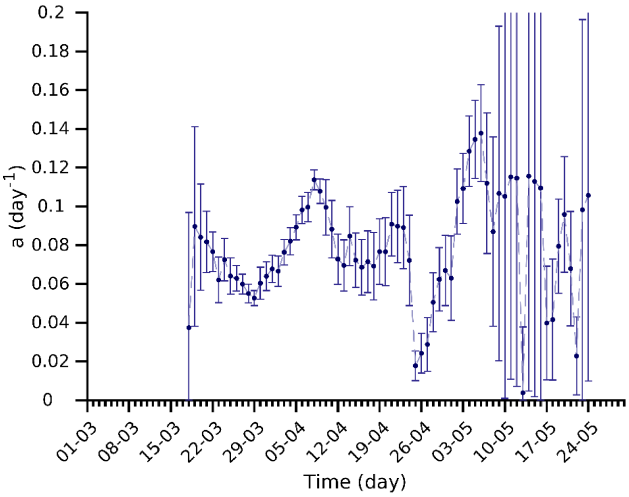


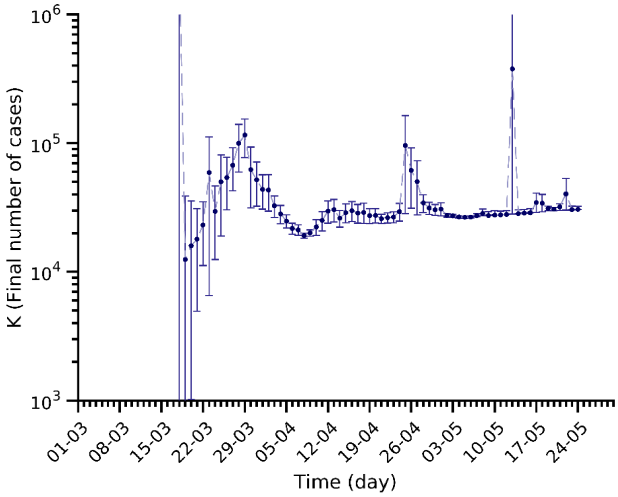


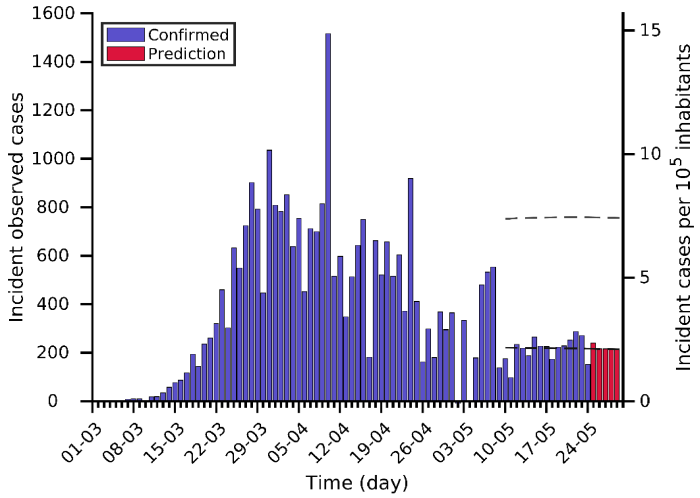
Portugal 24-05-2020. Population: 10.2M. Current cumulated incidence: 300/10^5

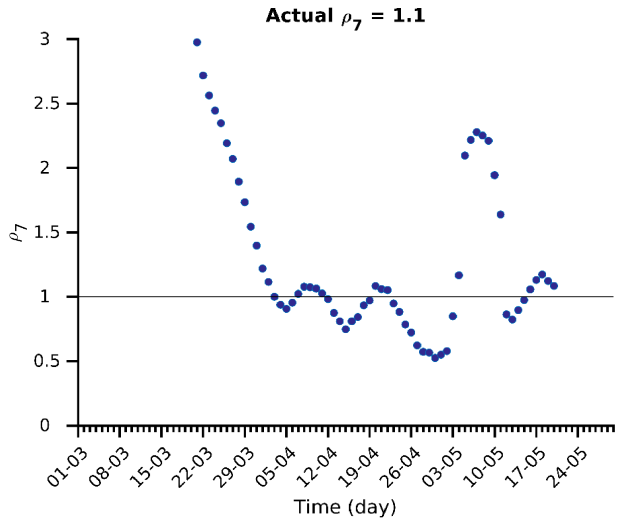


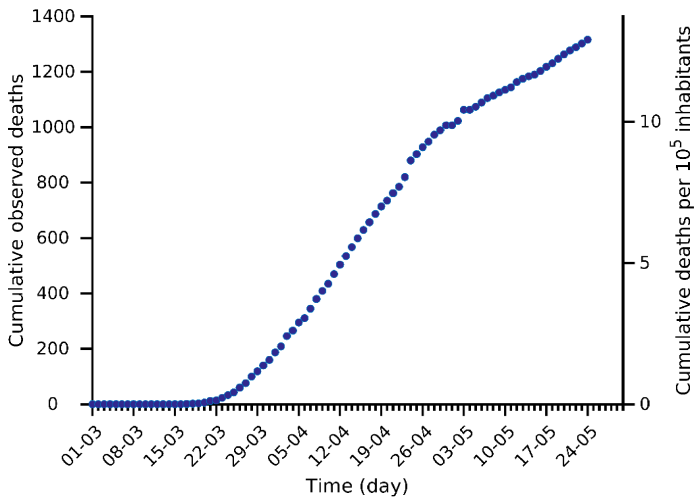


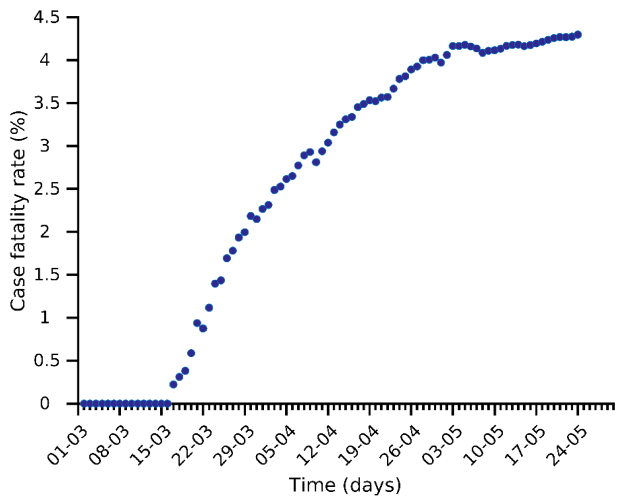




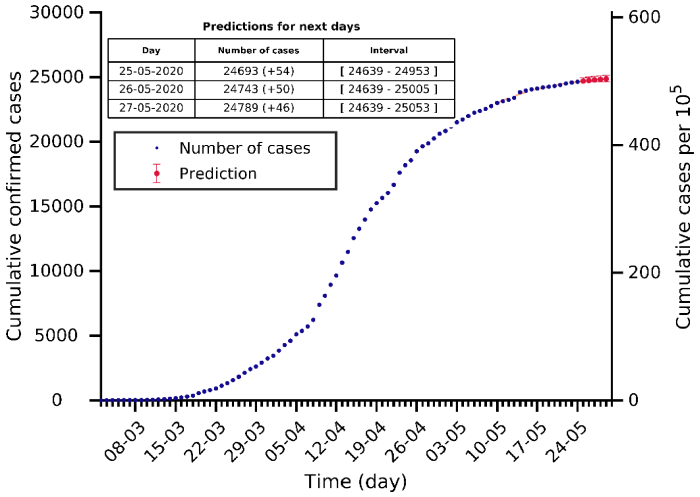


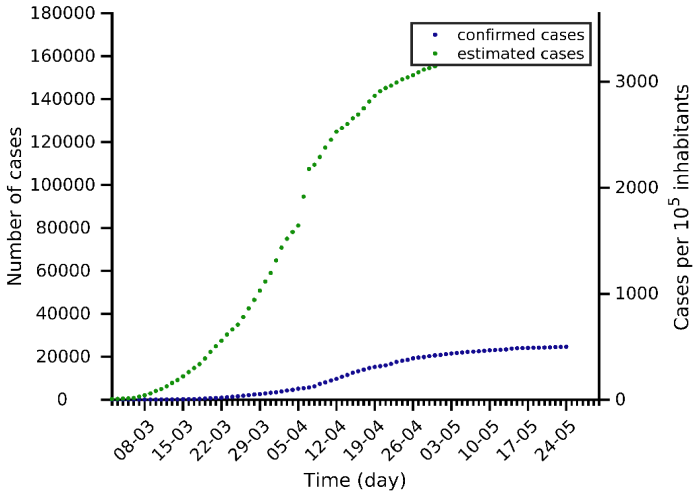


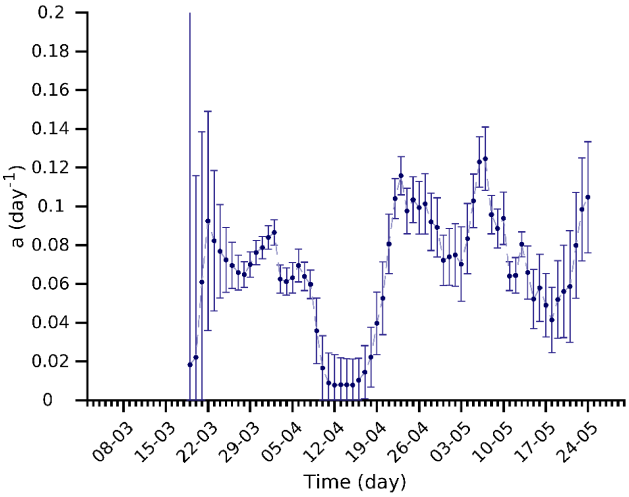


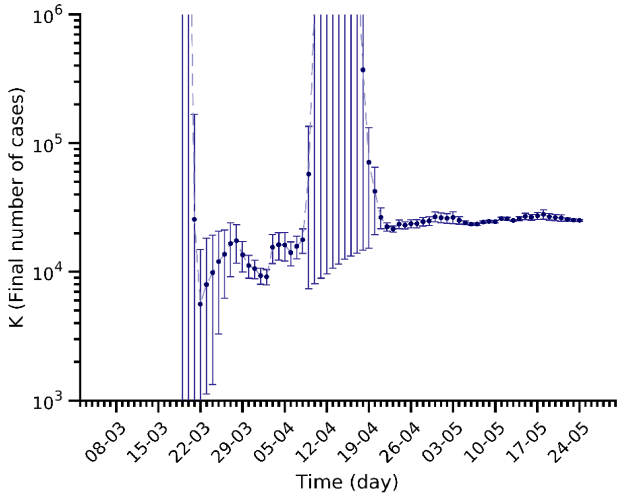


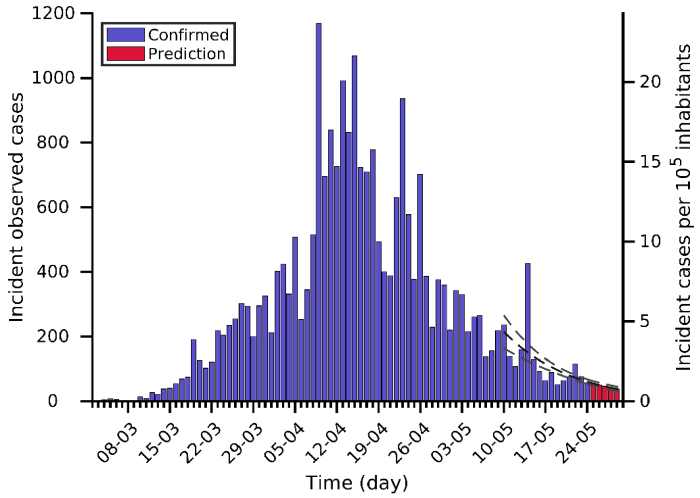
Ireland 24-05-2020. Population: 4.9M. Current cumulated incidence: 499/10^5

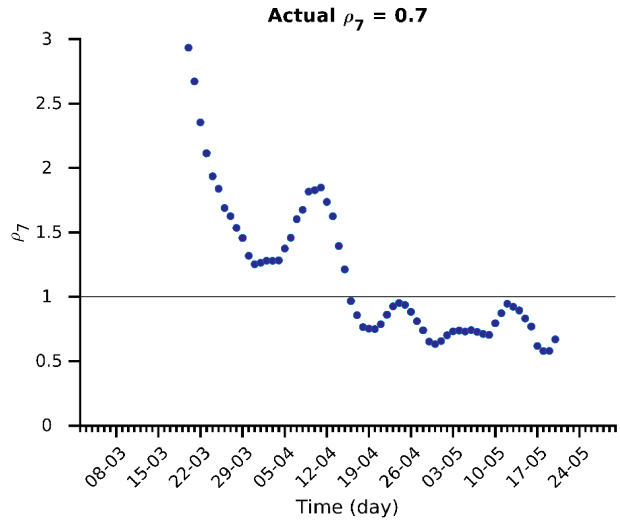


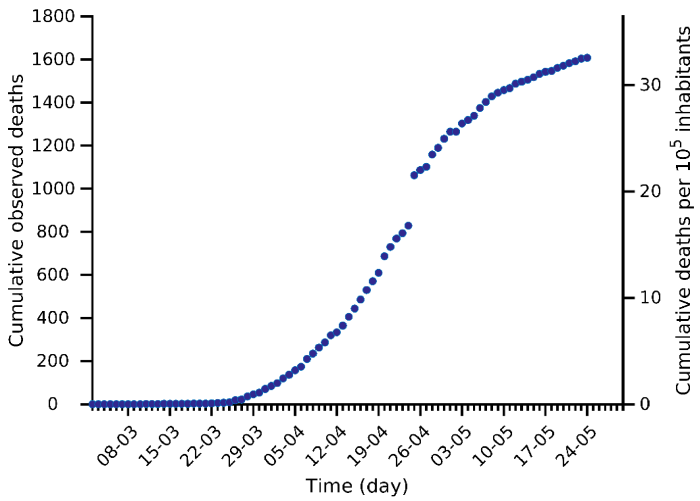


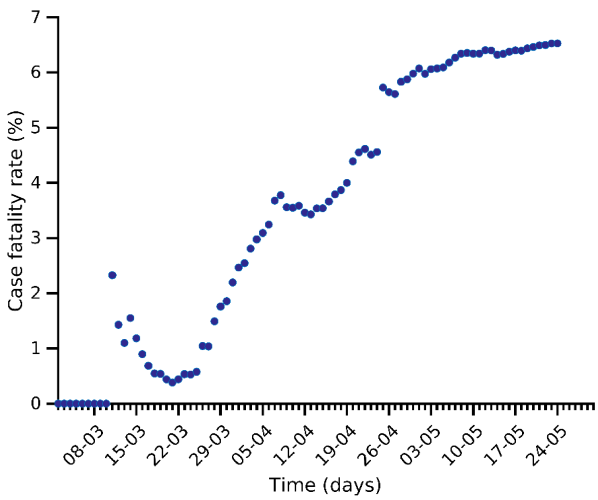




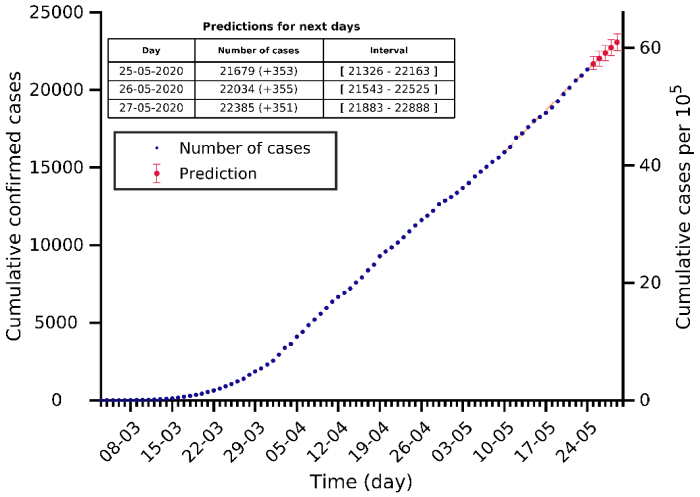


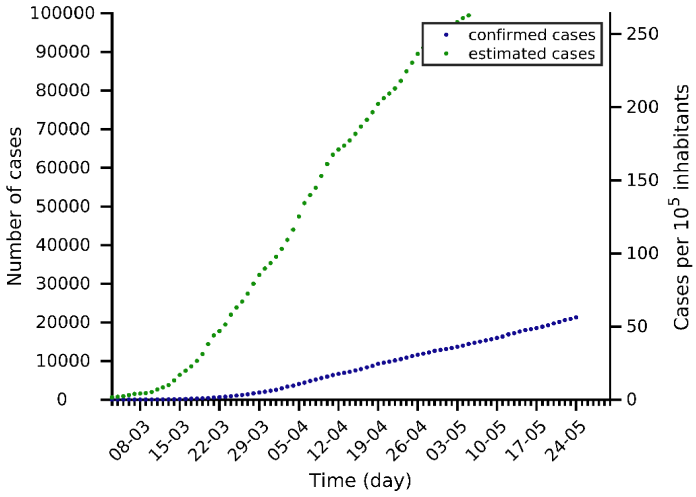


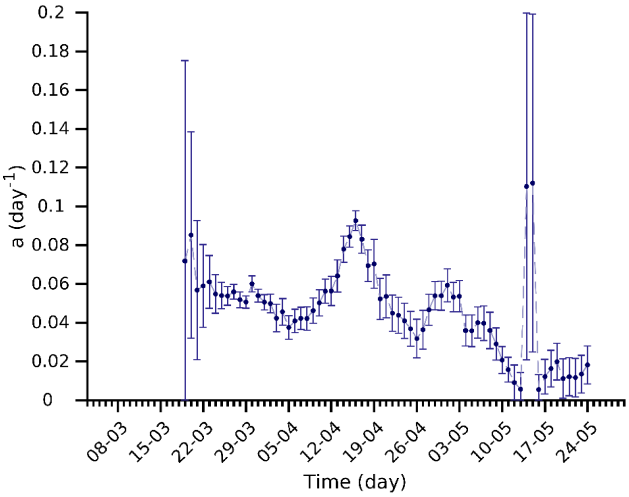


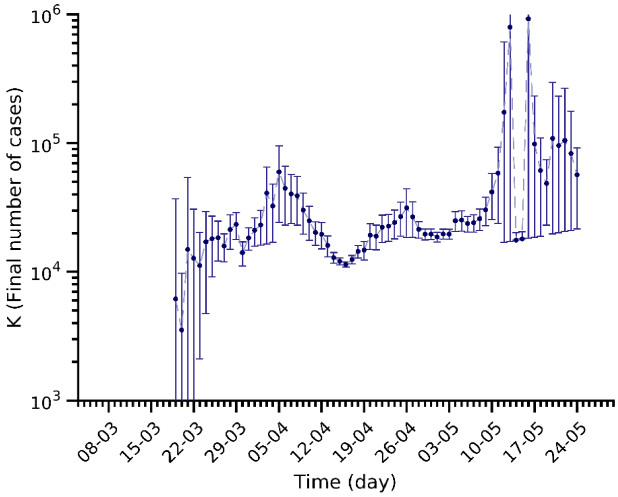


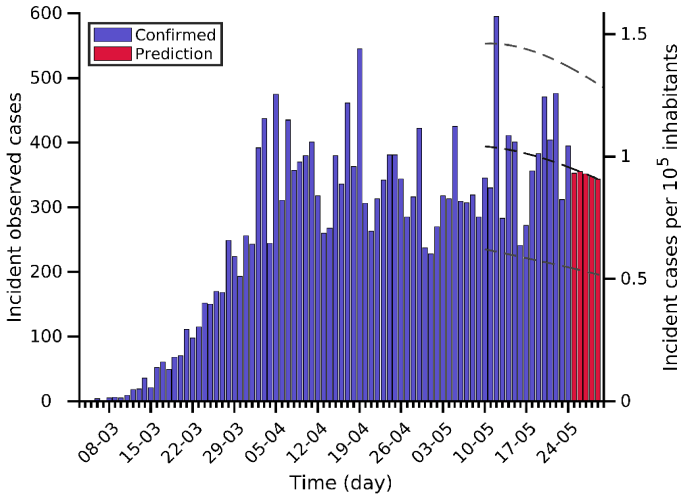
Poland 24-05-2020. Population: 37.8M. Current cumulated incidence: $56/10^5$

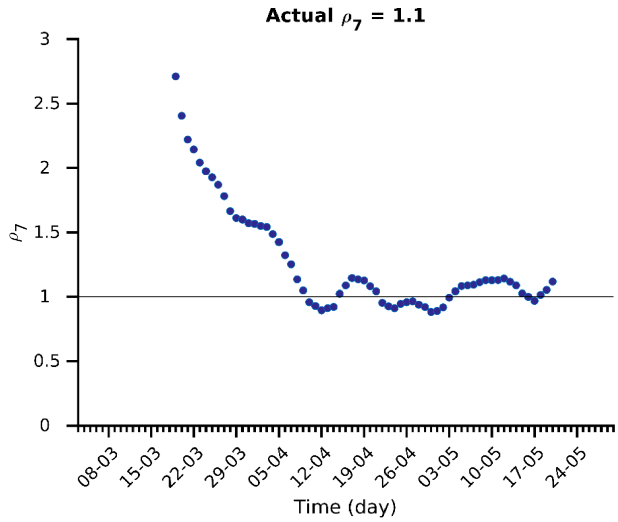


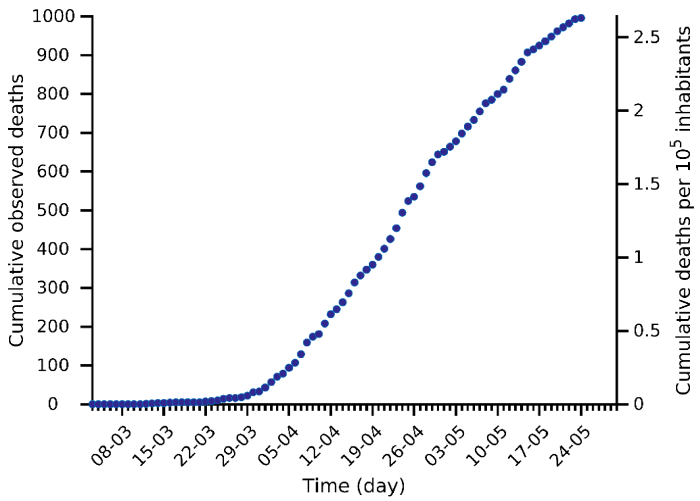


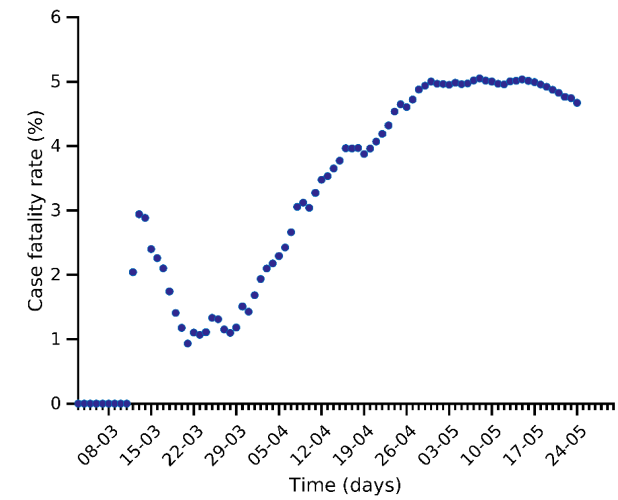




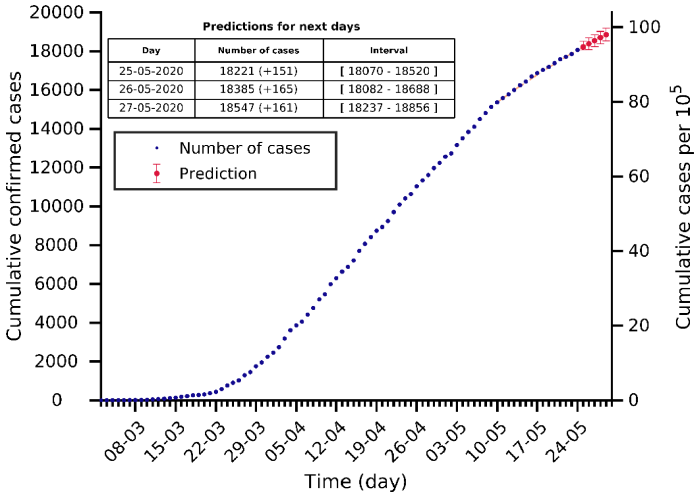


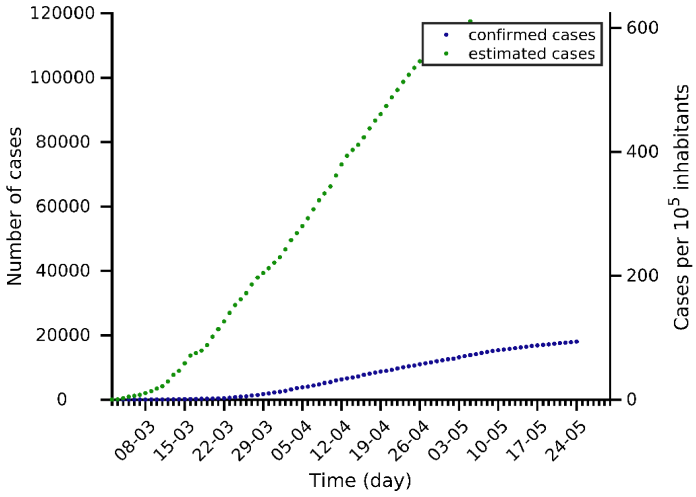


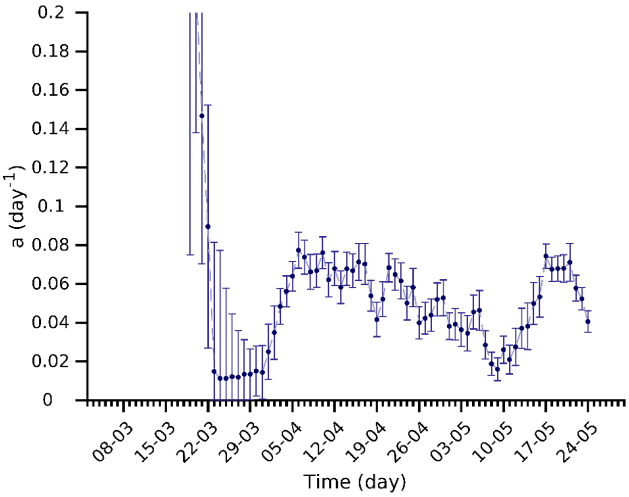


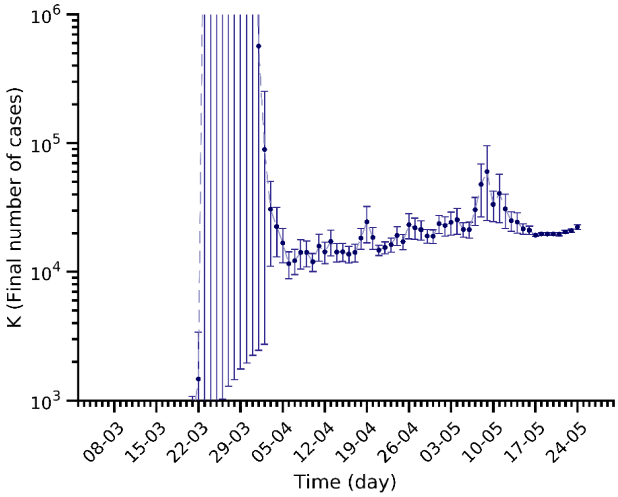


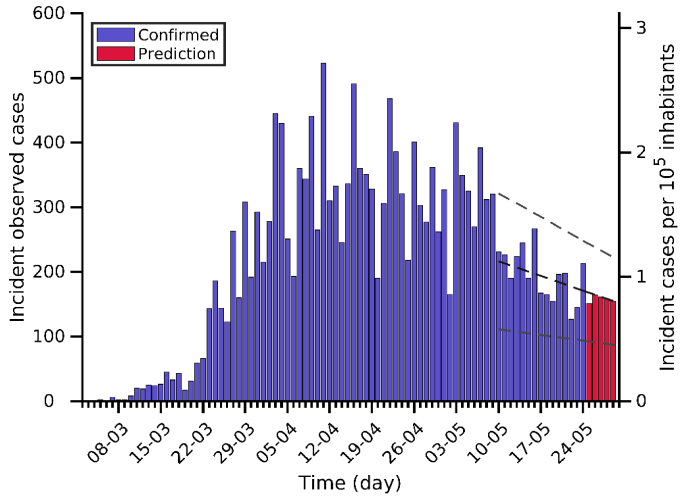
Romania 24-05-2020. Population: 19.2M. Current cumulated incidence: 94/10^5

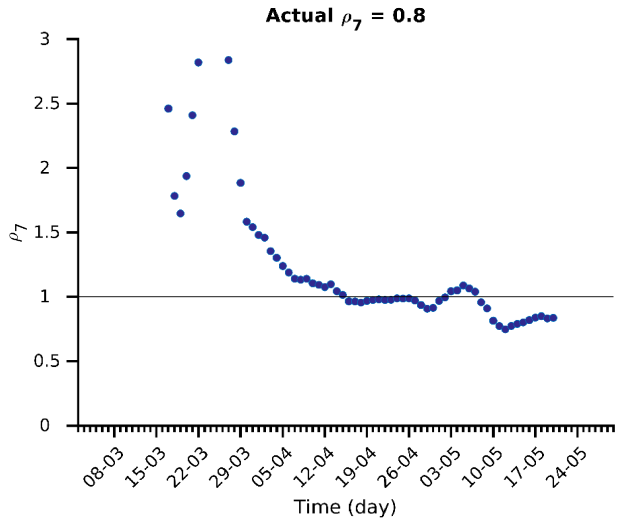


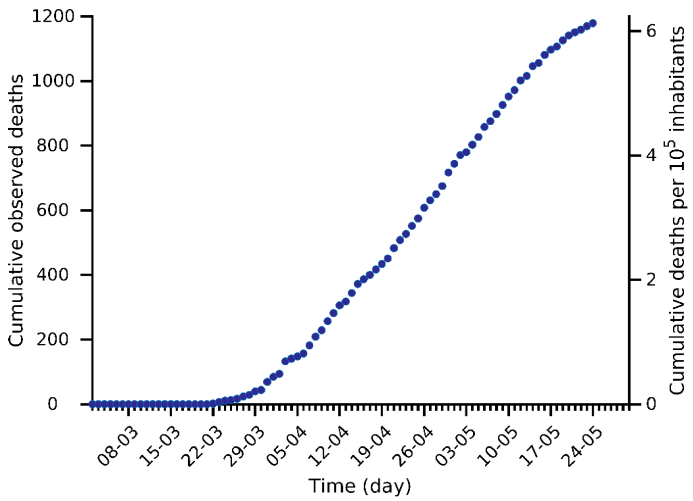


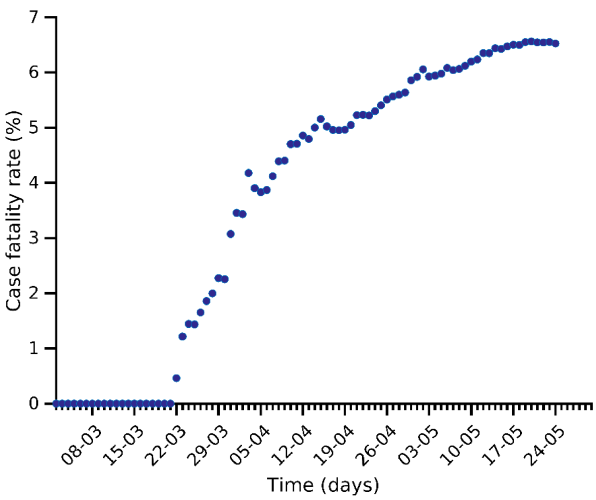




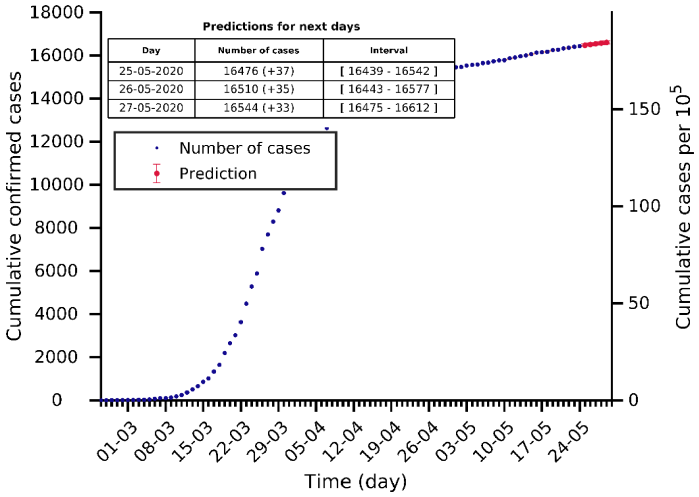


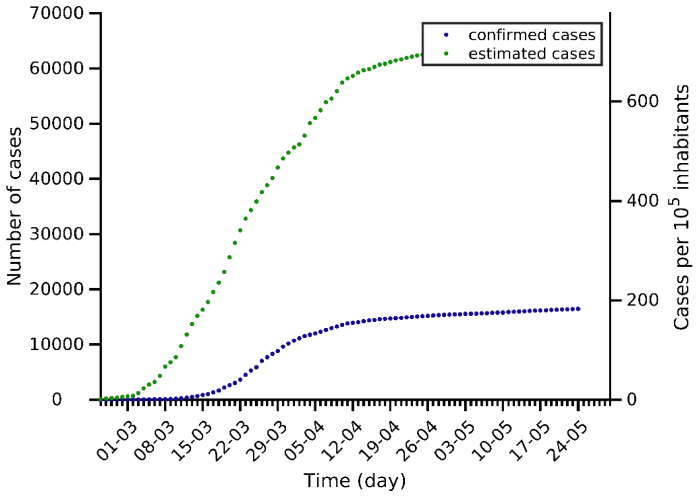


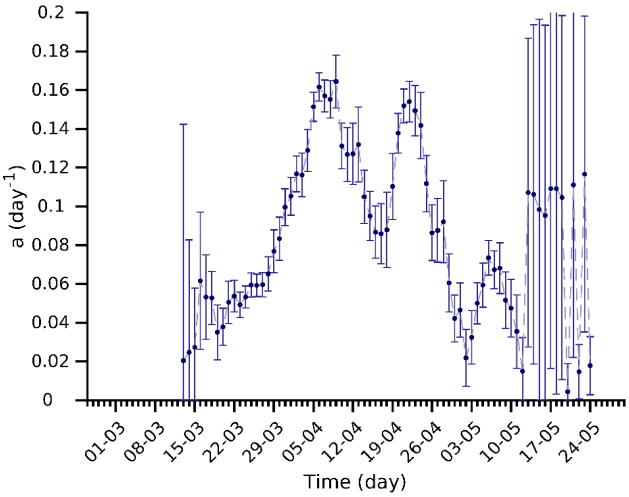


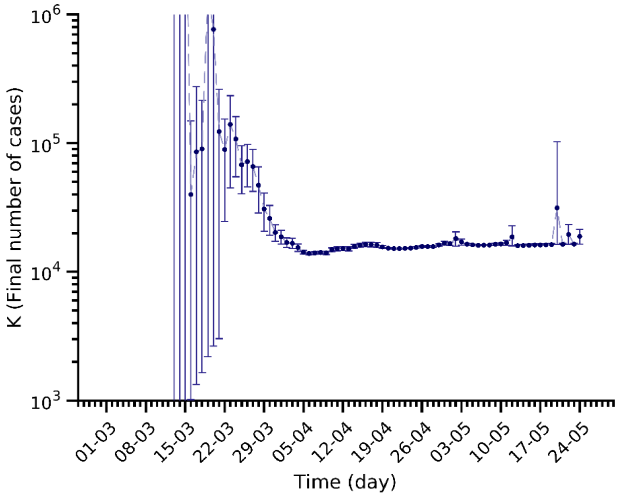


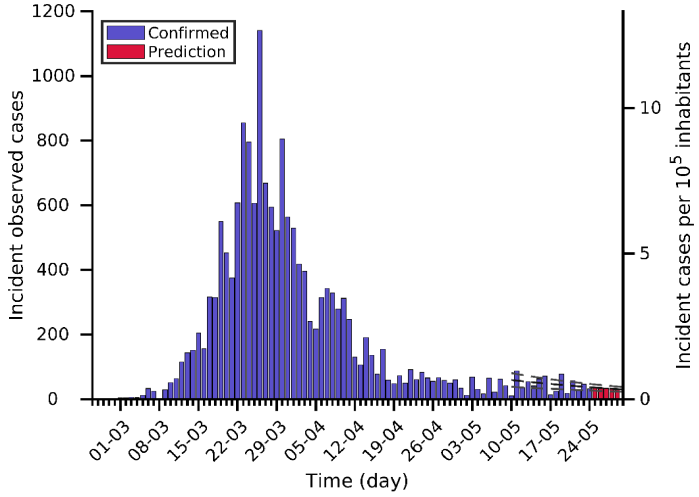
Austria 24-05-2020. Population: 9.0M. Current cumulated incidence: 183/10^5

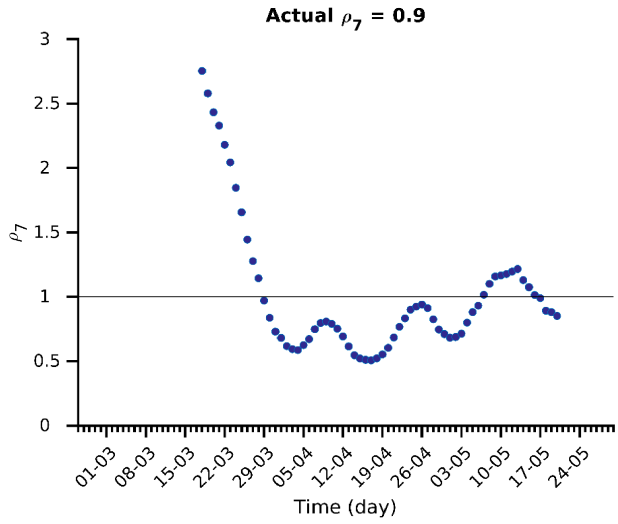


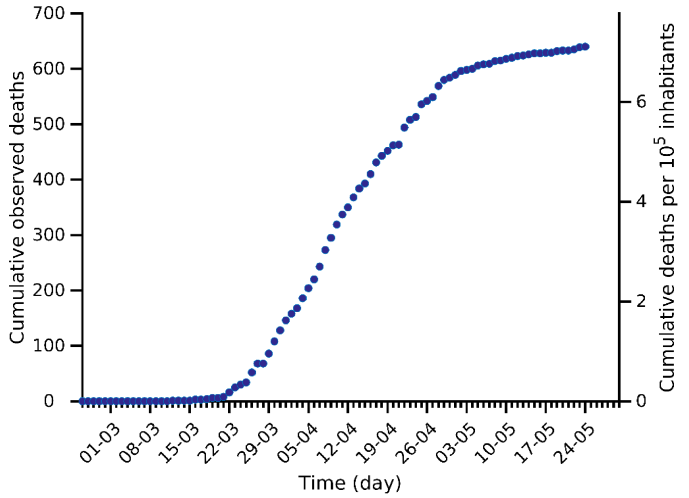


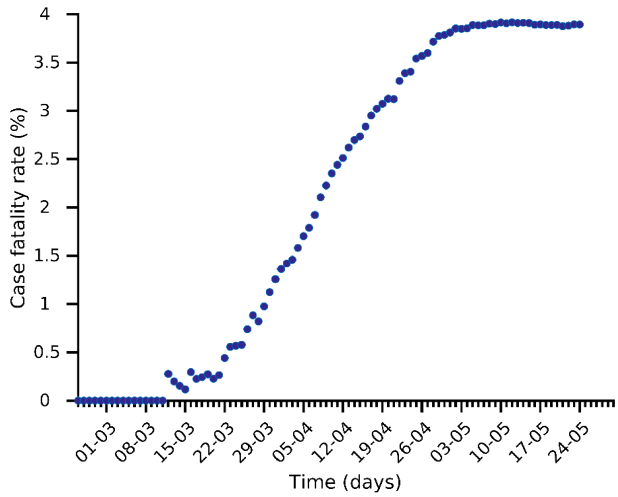




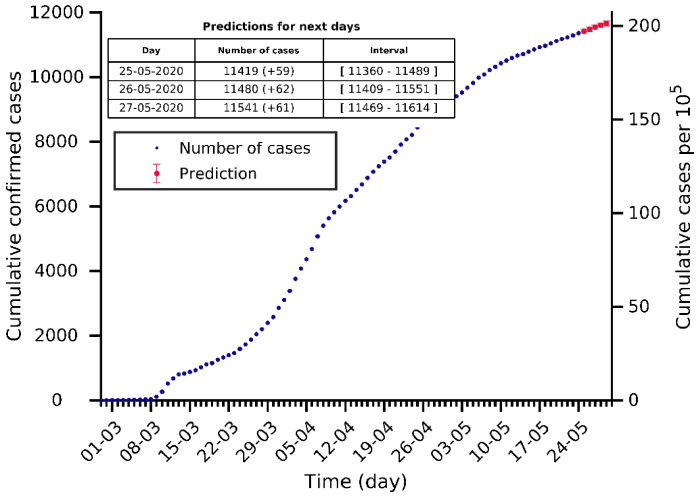


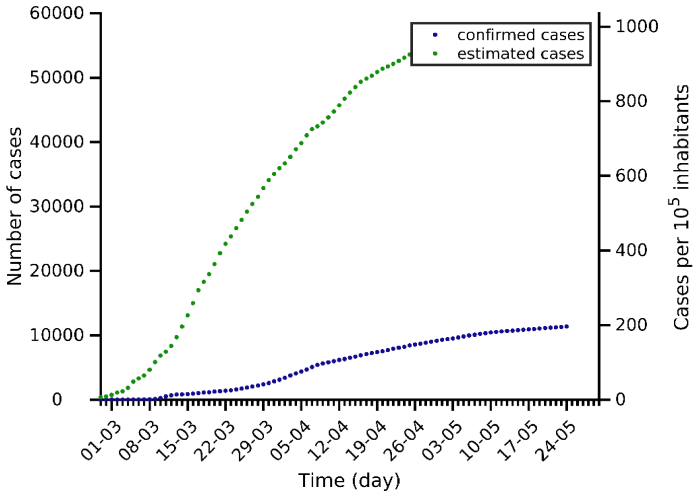


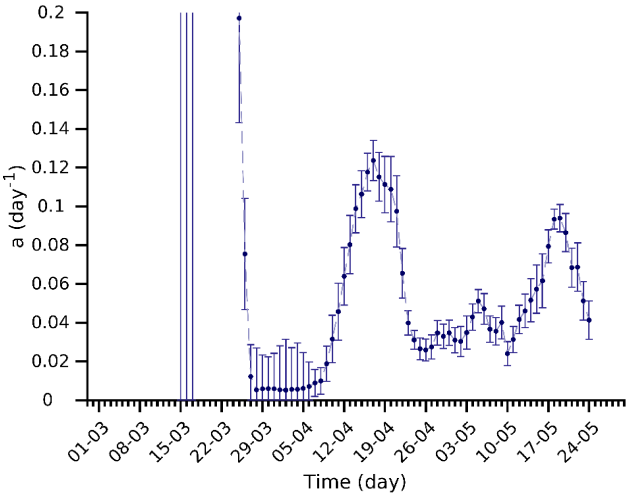


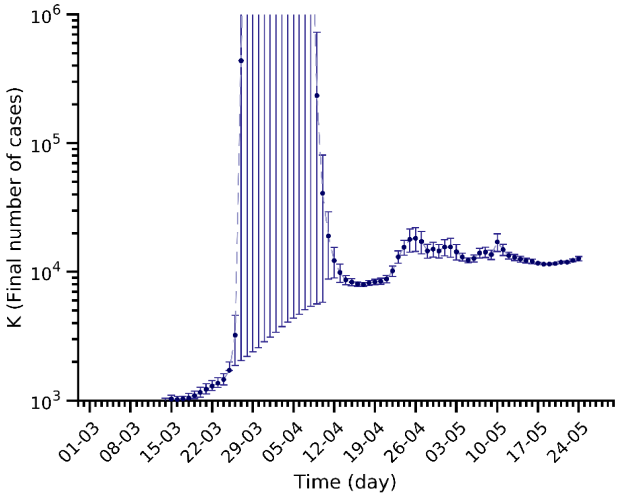


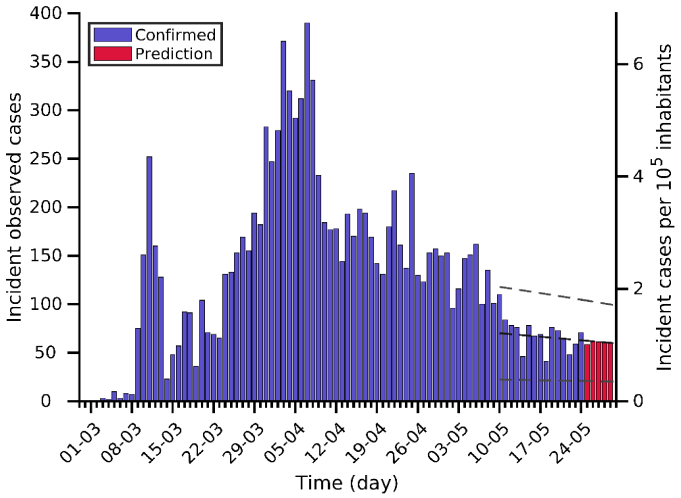
Denmark 24-05-2020. Population: 5.8M. Current cumulated incidence: 196/10⁵

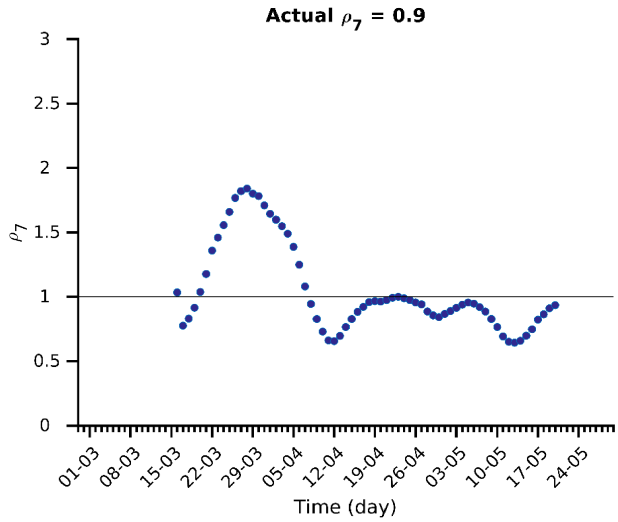


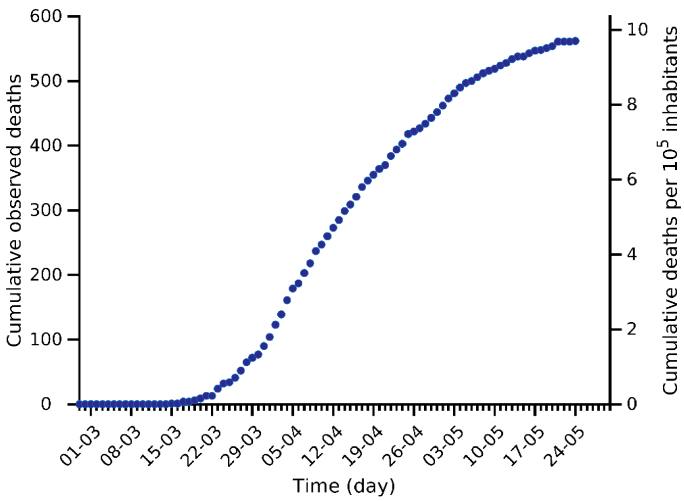


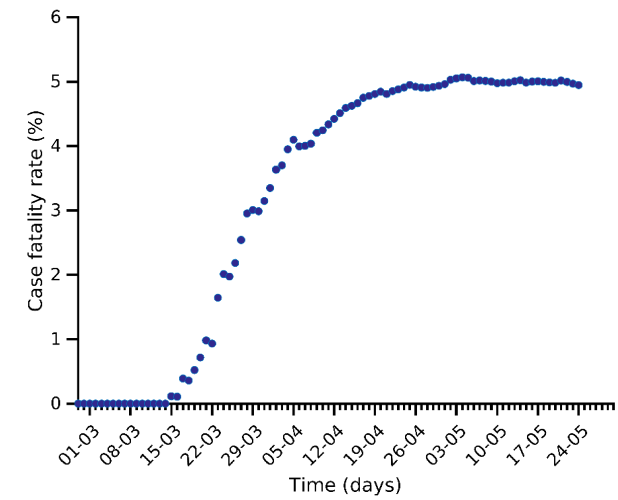




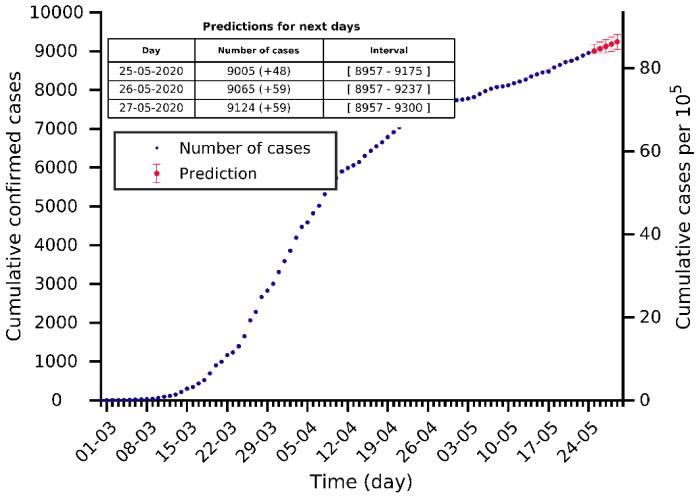


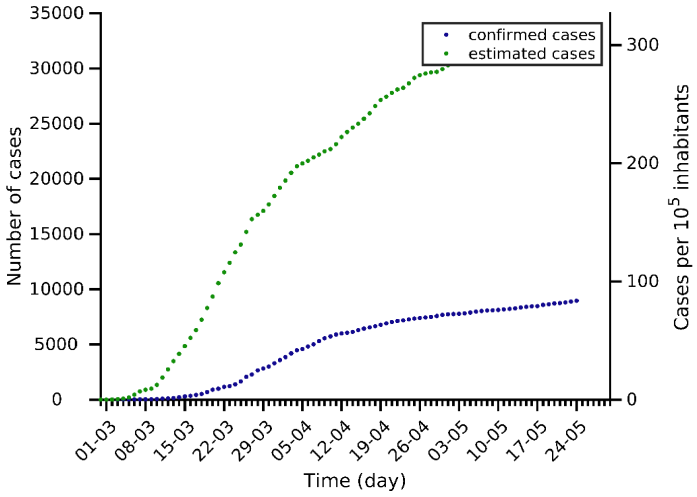


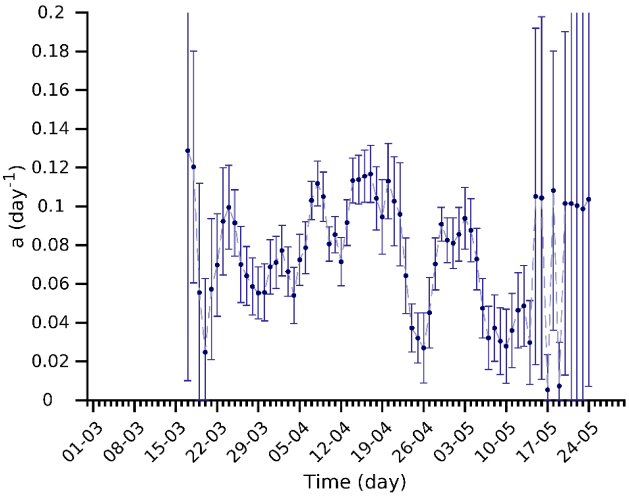


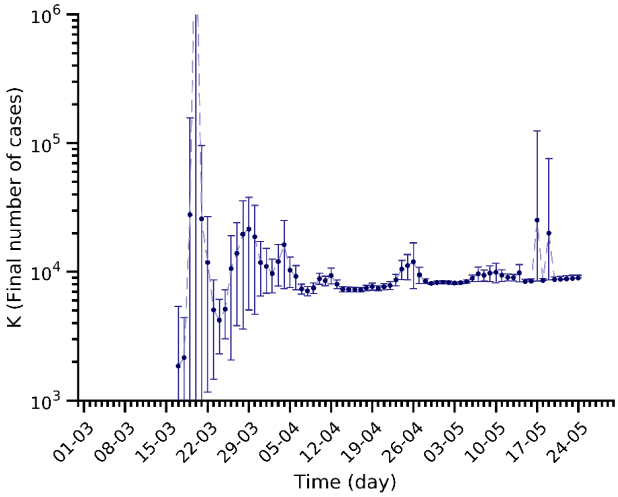


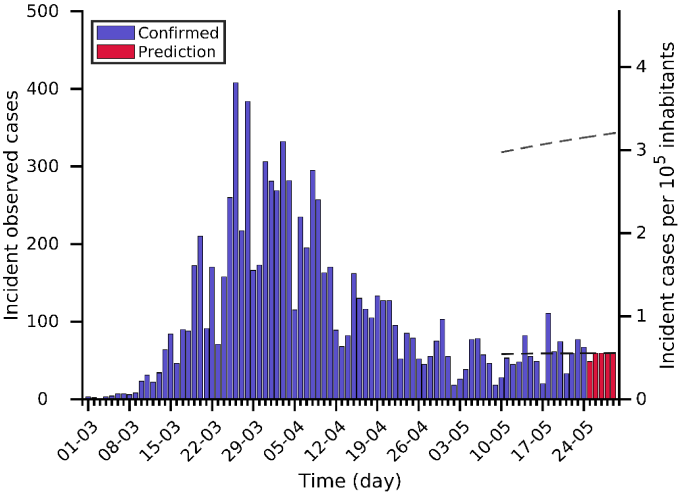
Czech Rep 24-05-2020. Population: 10.7M. Current cumulated incidence: 84/10^5

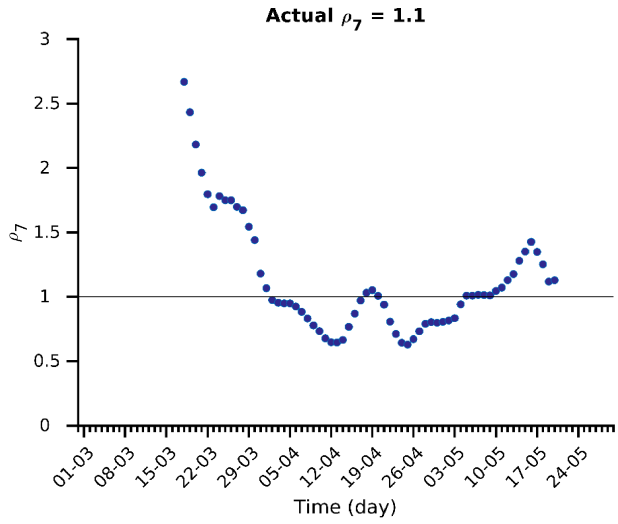


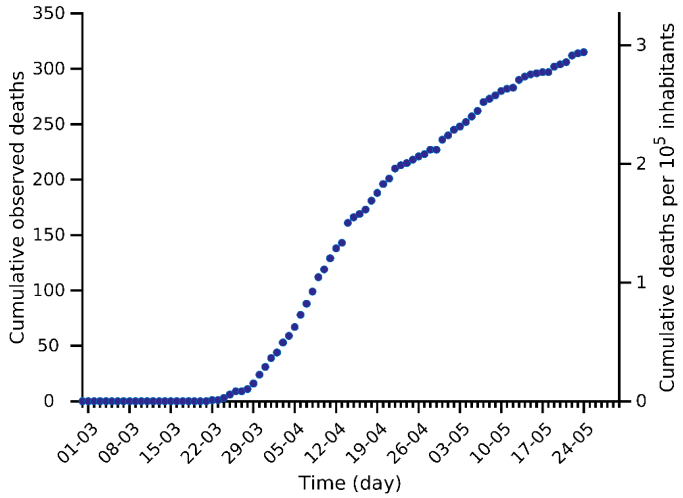


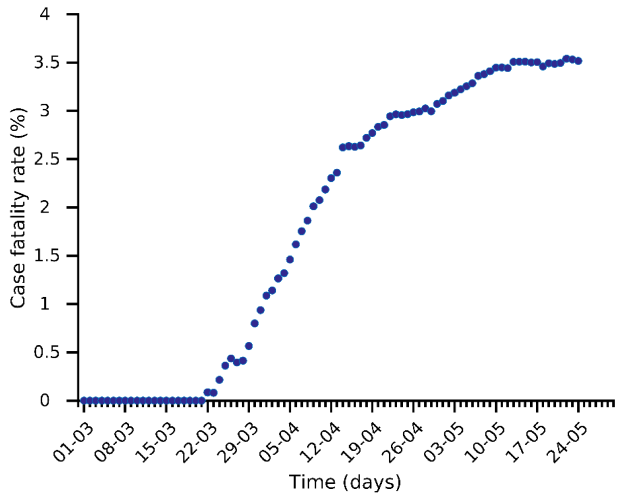




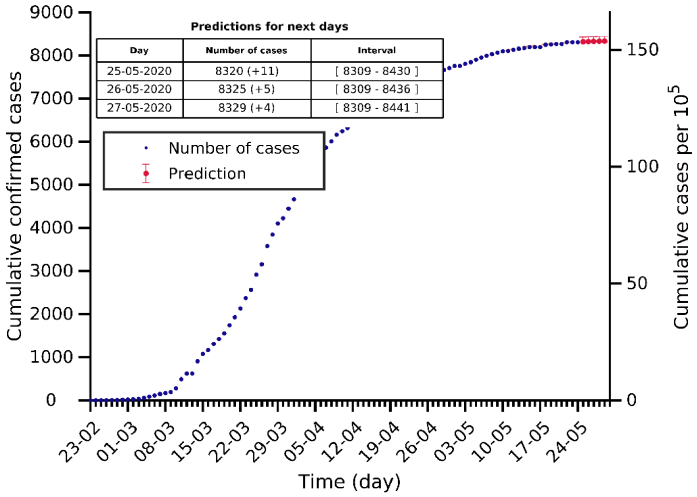


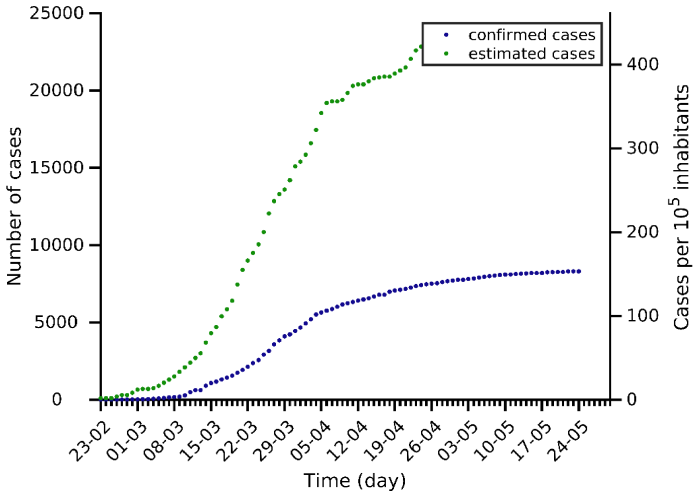


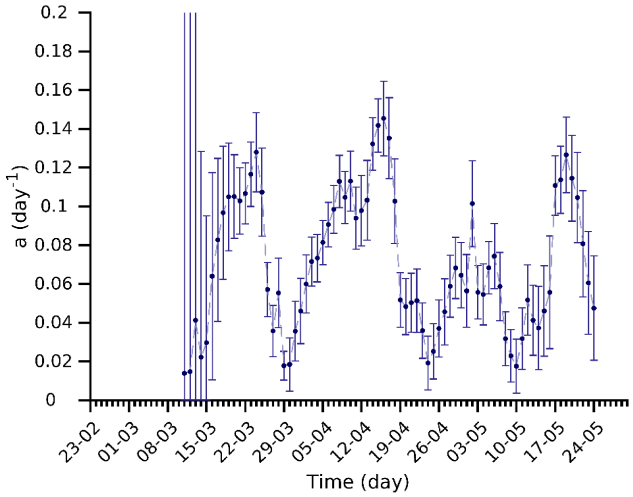


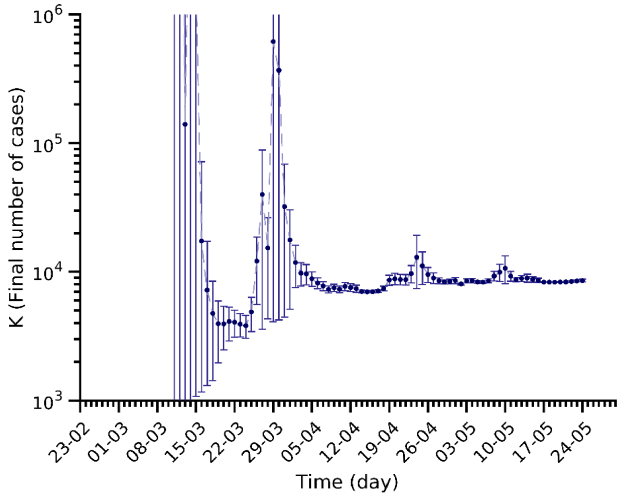


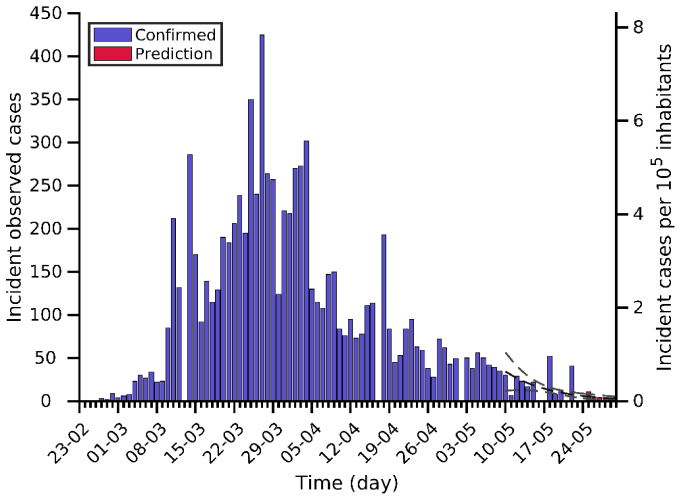
Norway 24-05-2020. Population: 5.4M. Current cumulated incidence: 153/10⁵

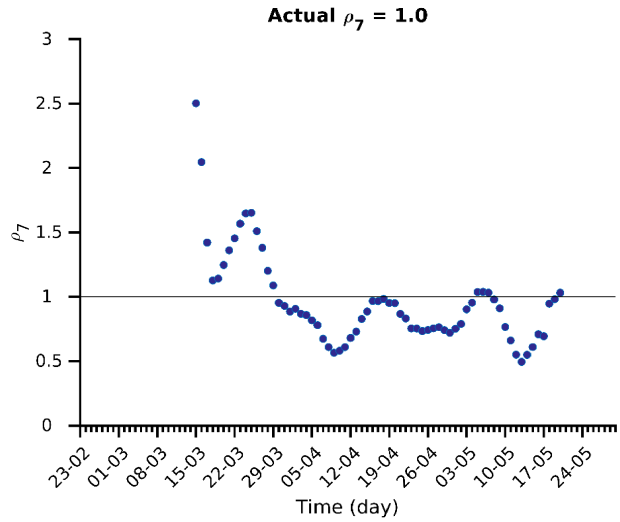


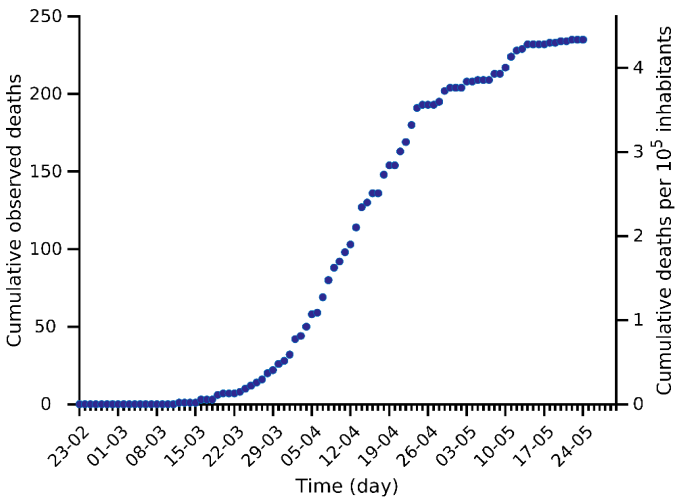


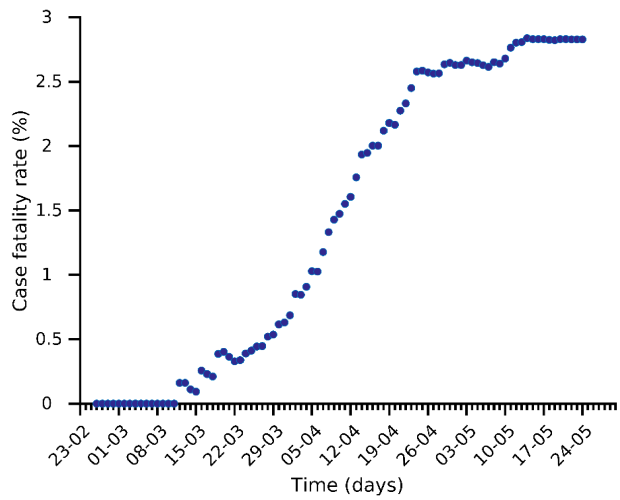




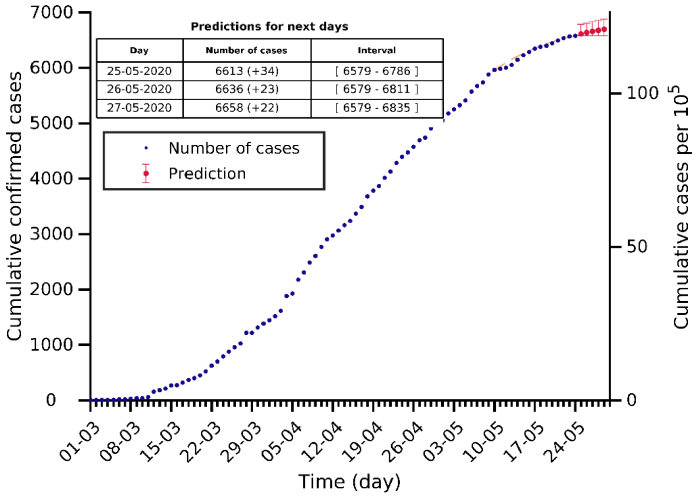


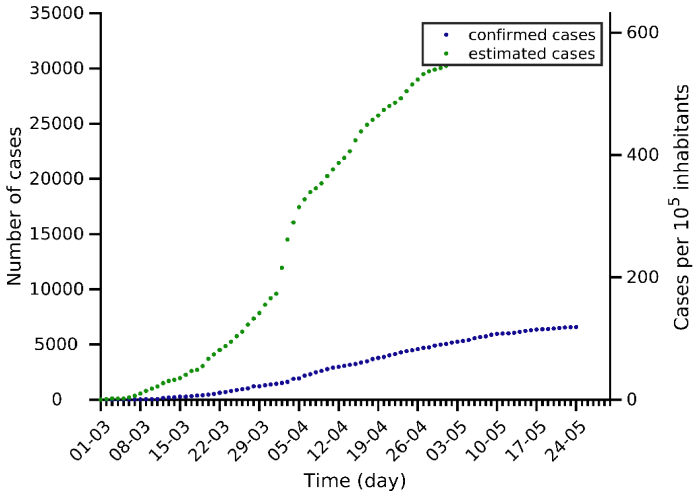


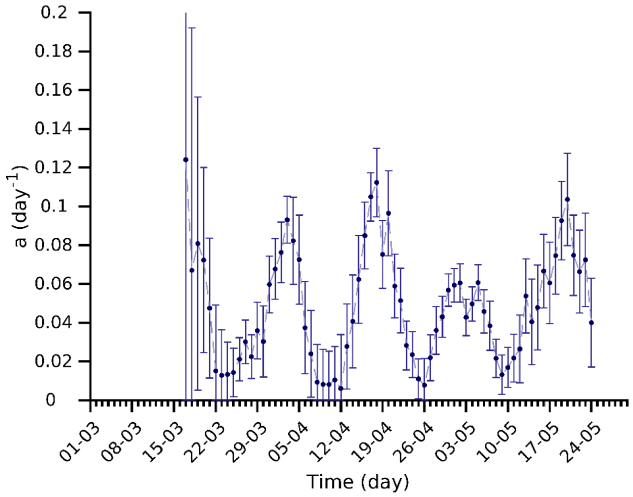


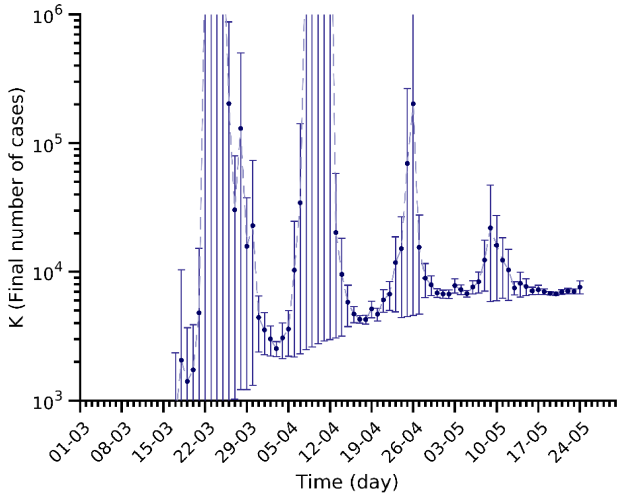


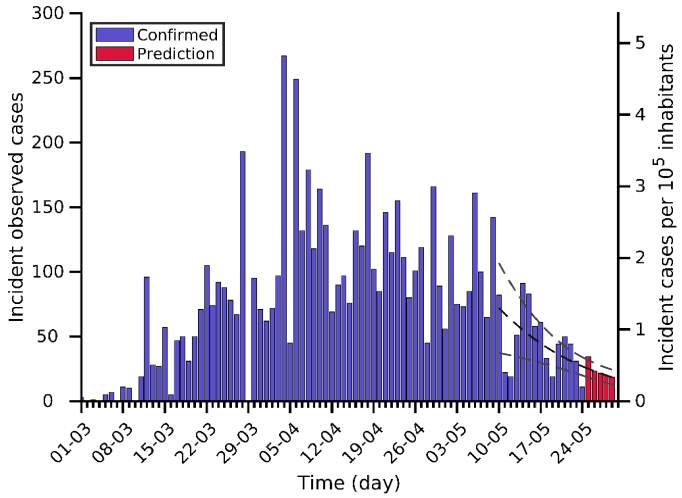
Finland 24-05-2020. Population: 5.5M. Current cumulated incidence: 119/10⁵

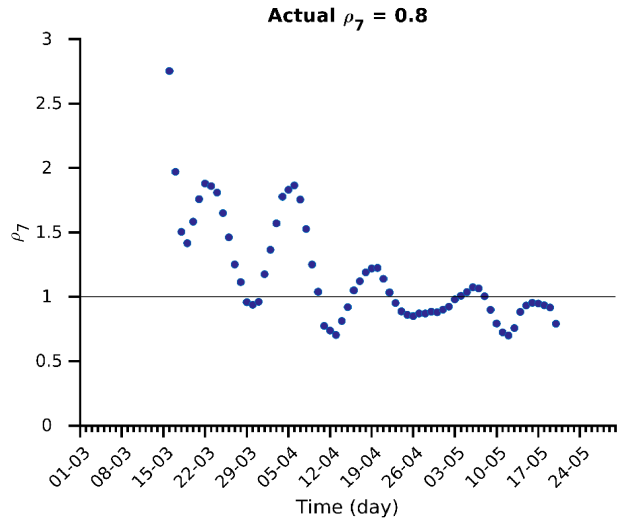


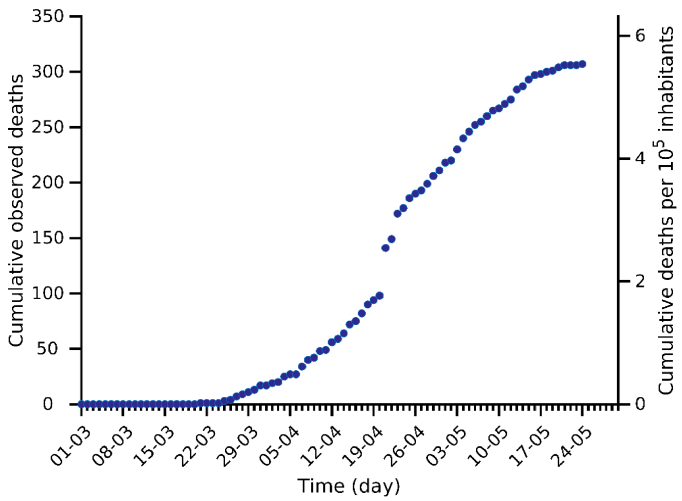


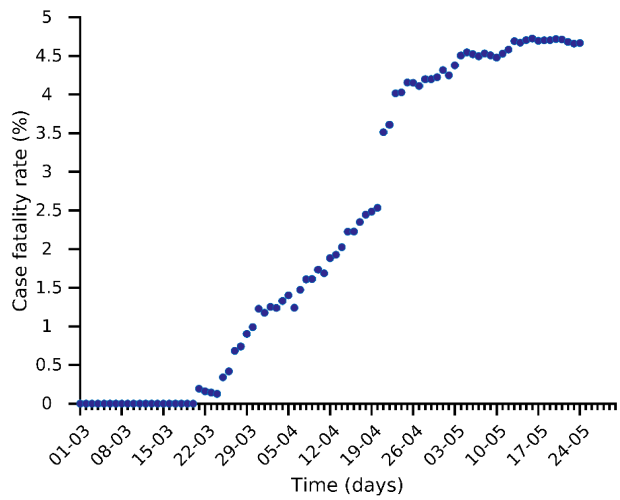




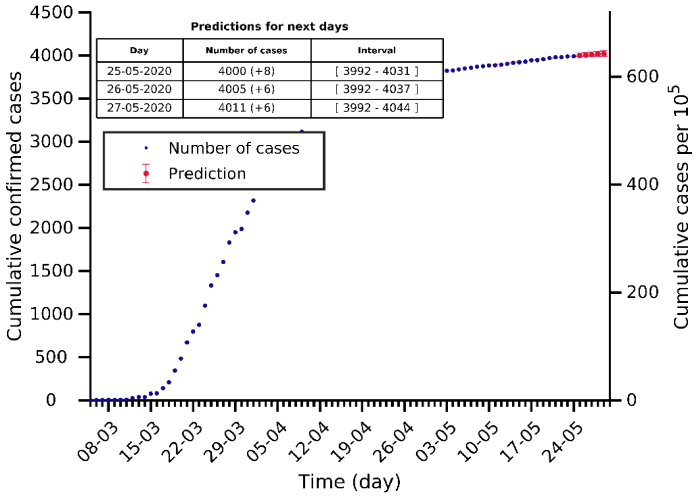


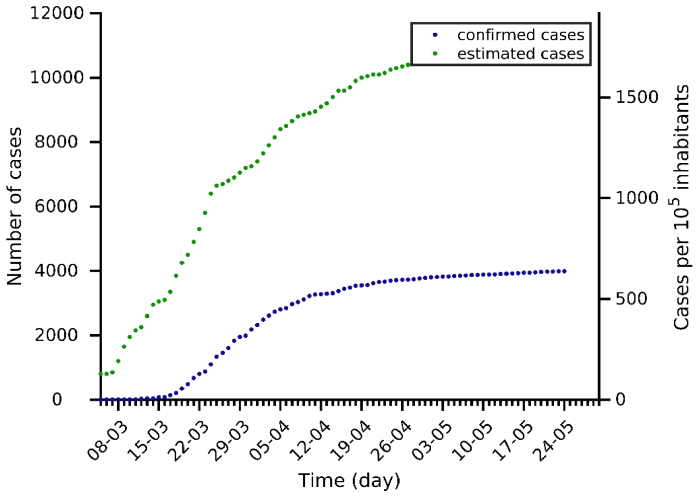


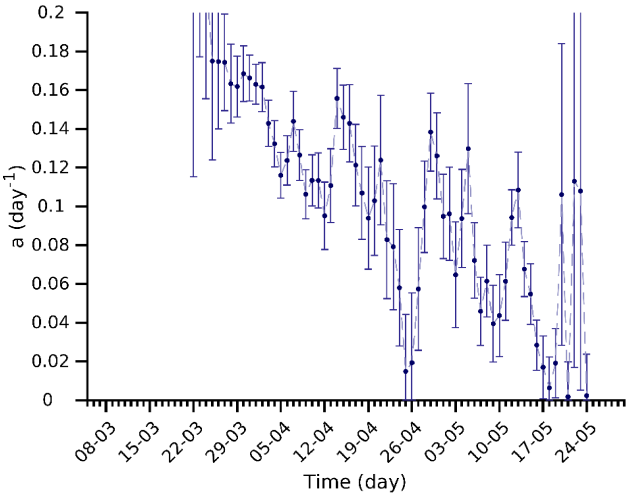


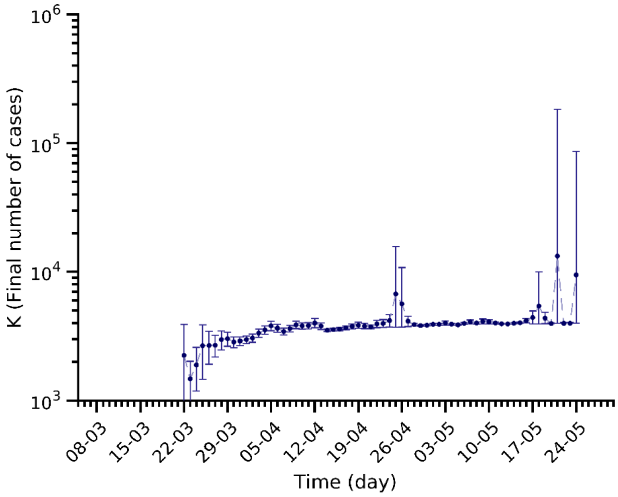


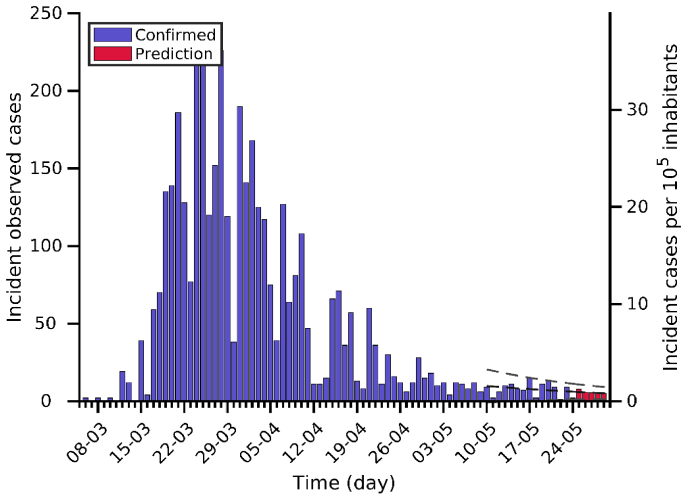
Luxembourg 24-05-2020. Population: 0.6M. Current cumulated incidence: 638/10^5

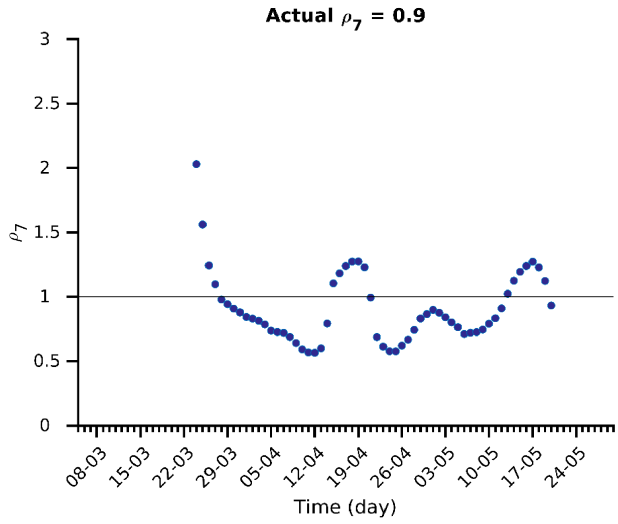


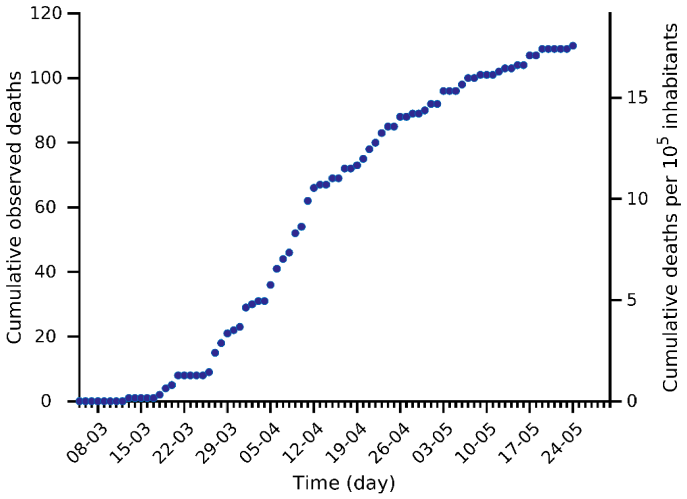


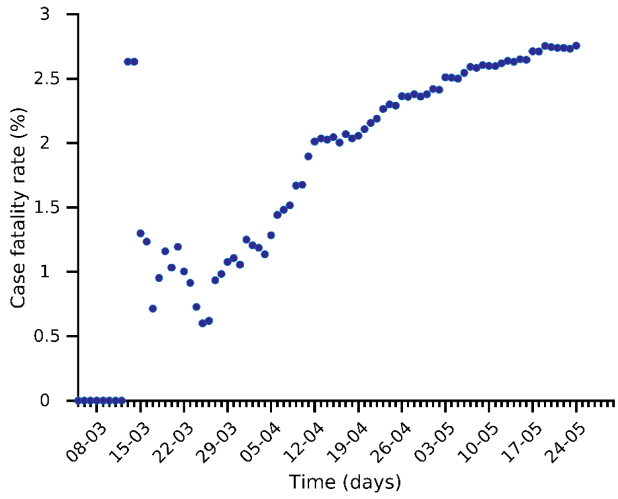




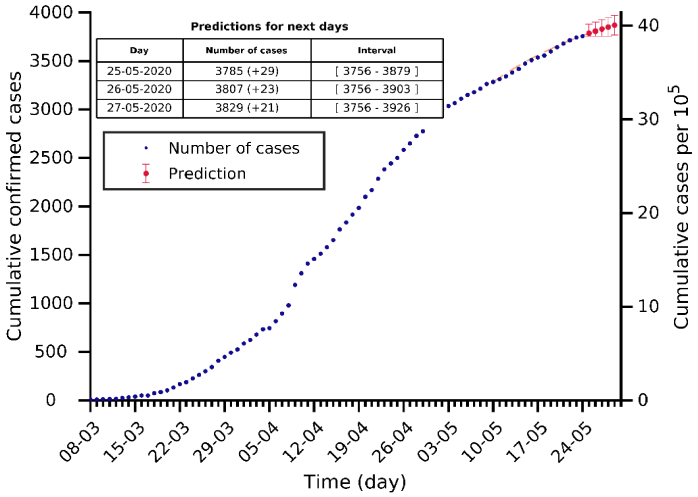


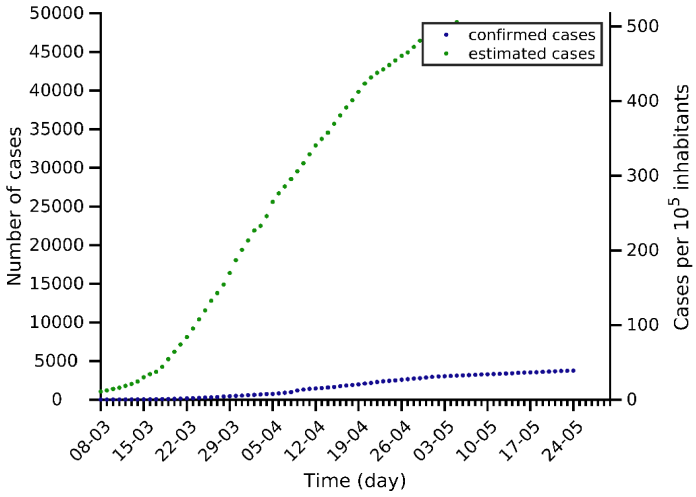


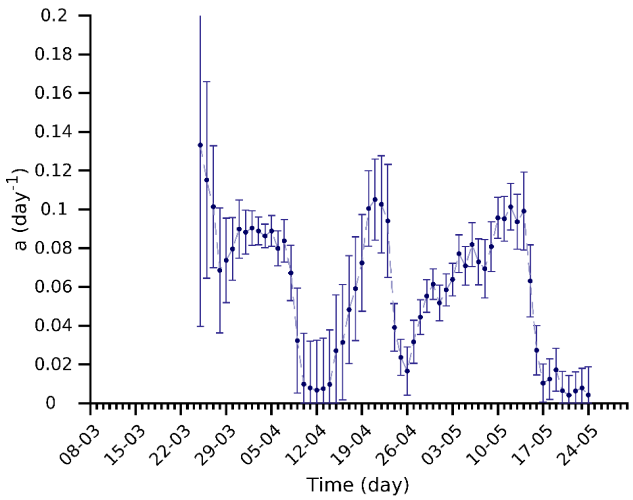


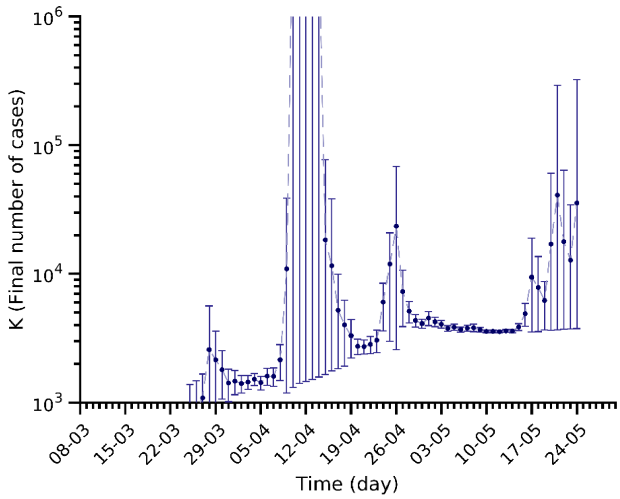


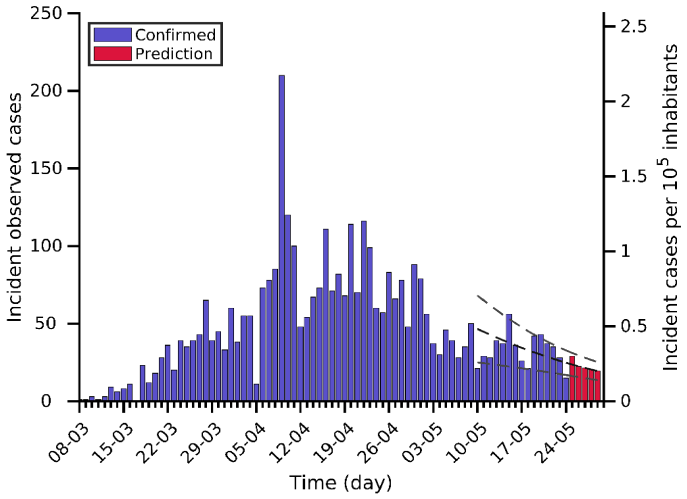
Hungary 24-05-2020. Population: 9.7M. Current cumulated incidence: $39/10^5$

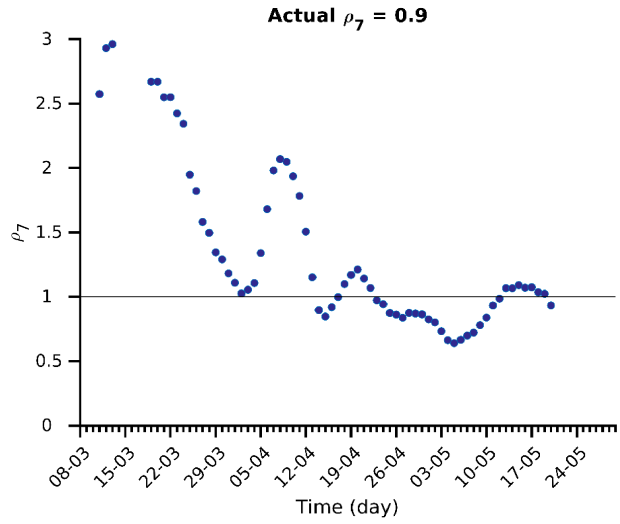


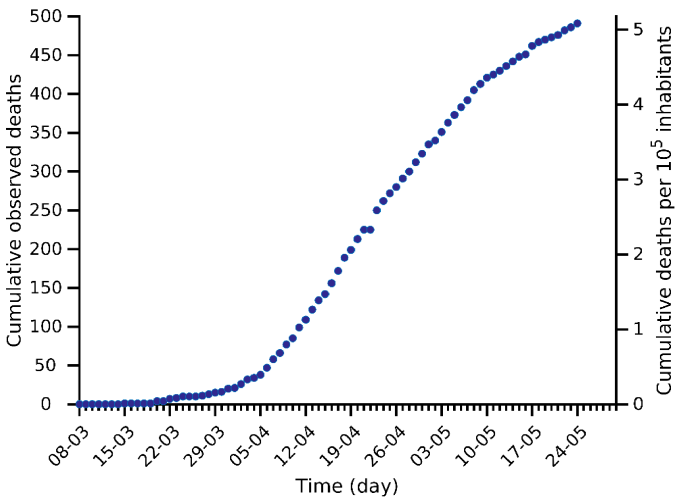


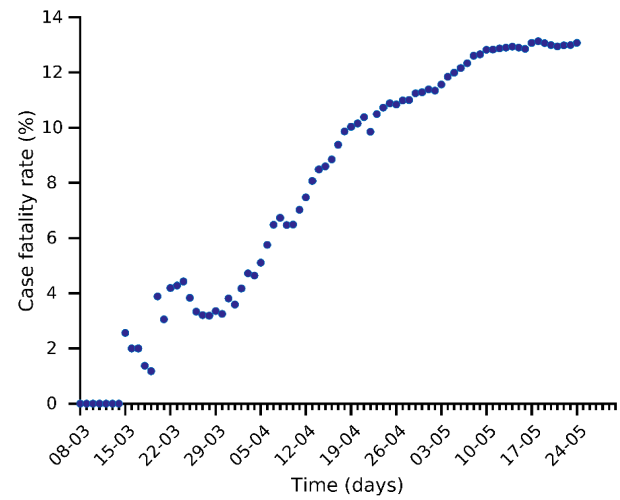




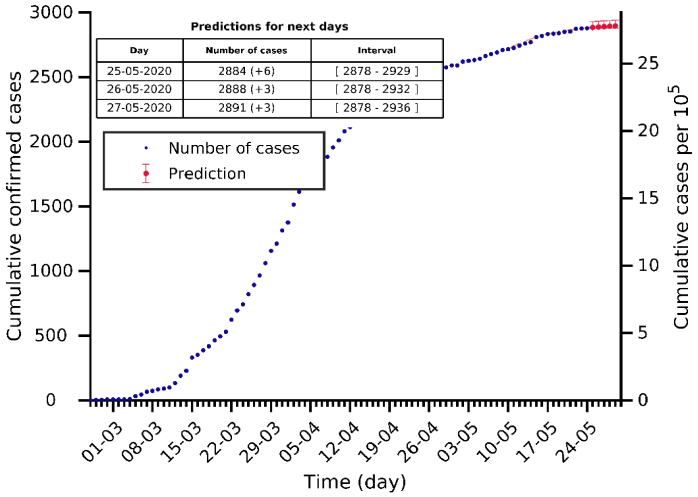


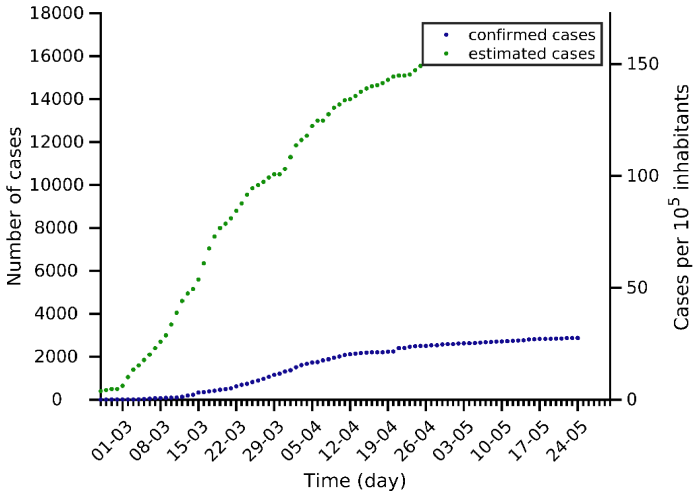


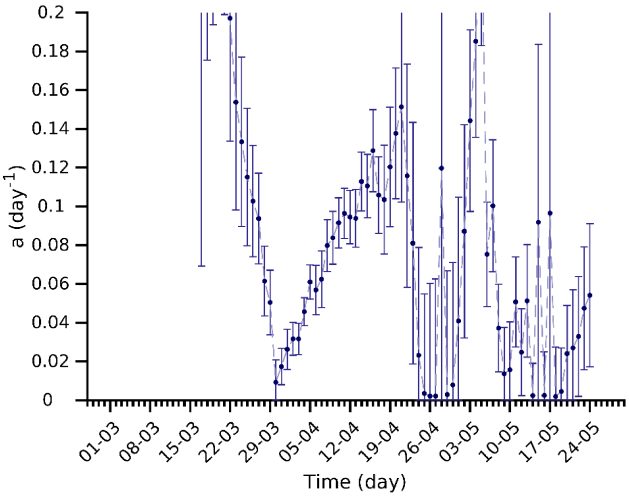


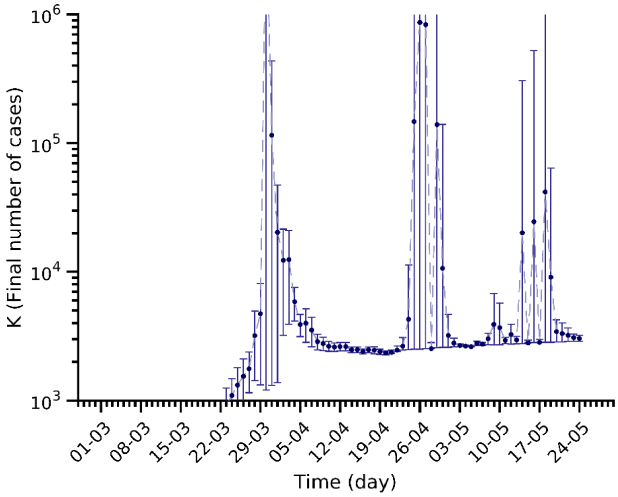


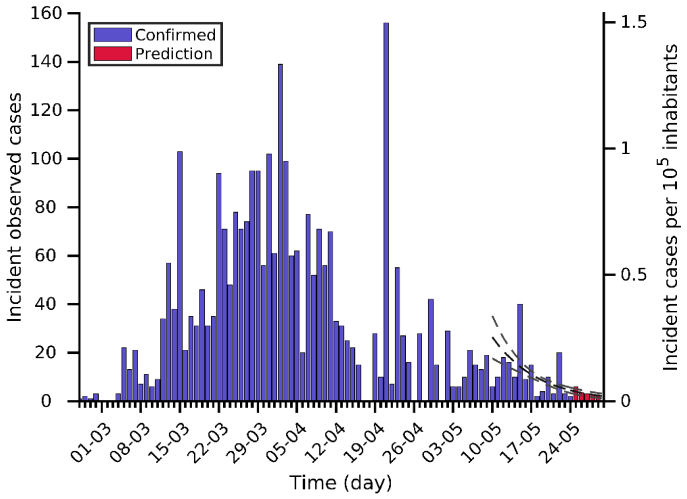
Greece 24-05-2020. Population: 10.4M. Current cumulated incidence: $28/10^5$

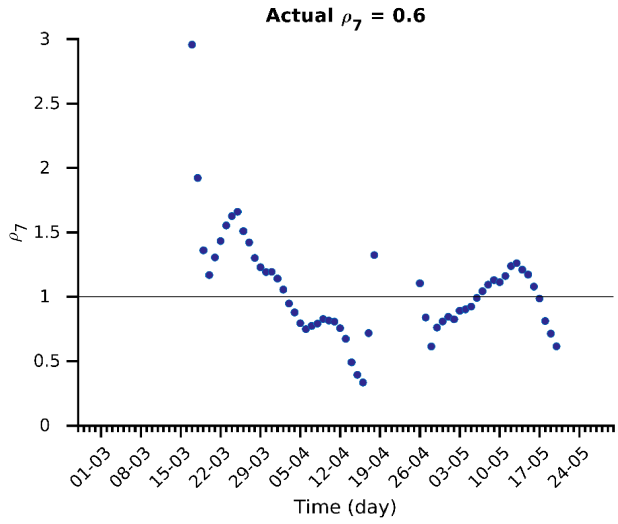


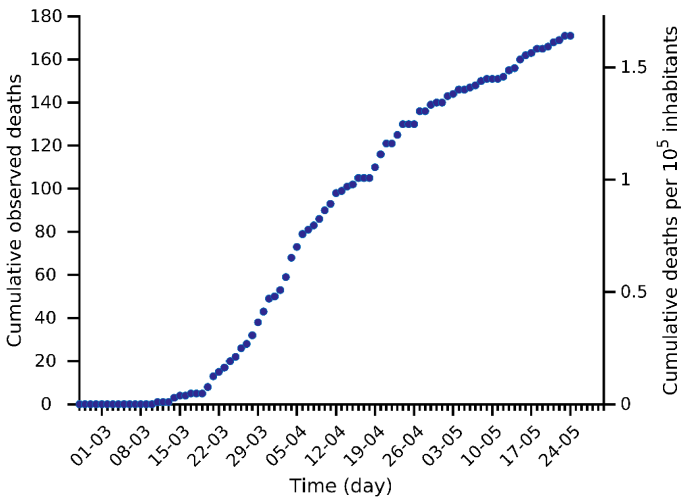


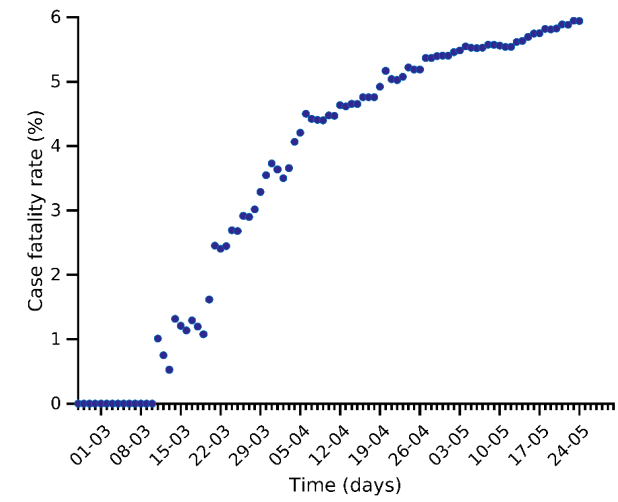




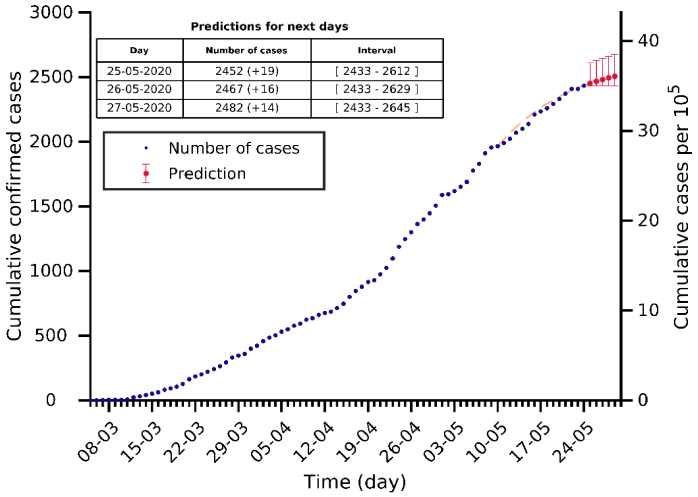


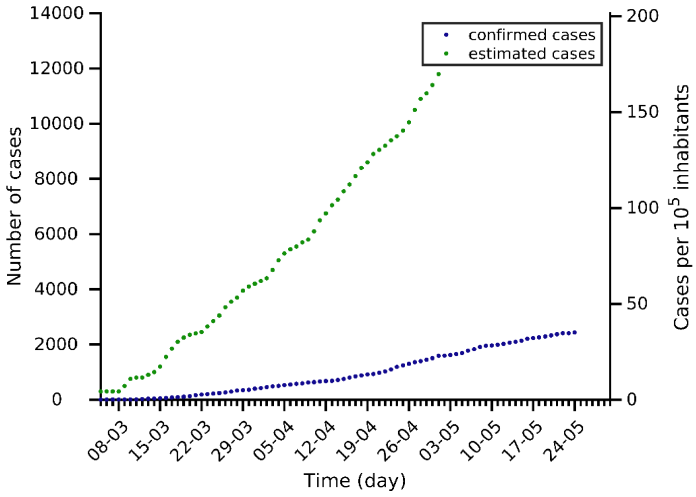


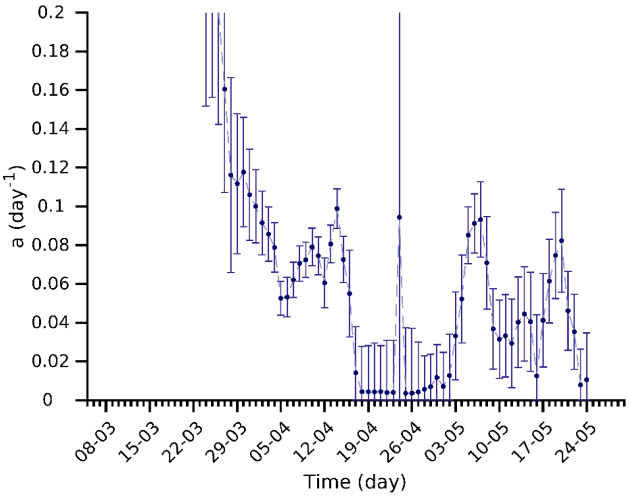


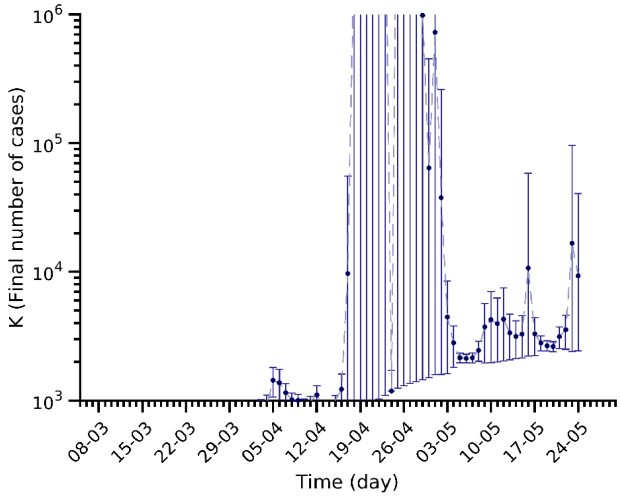


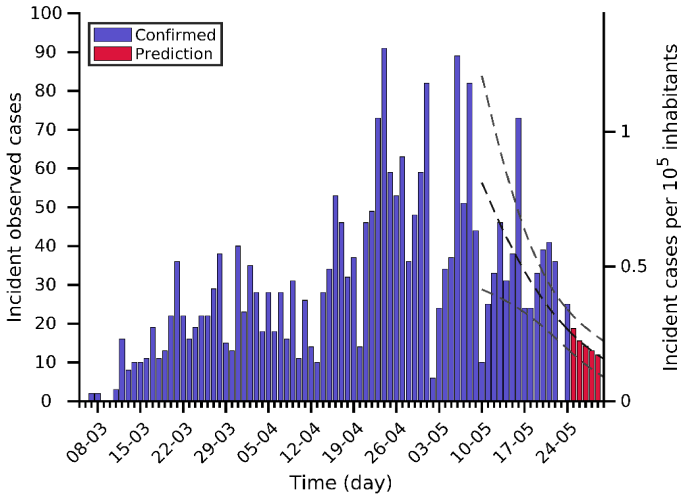
Bulgaria 24-05-2020. Population: 6.9M. Current cumulated incidence: $35/10^5$

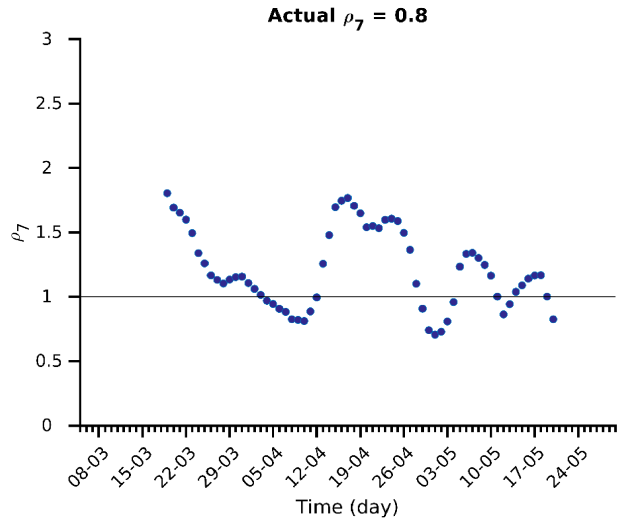


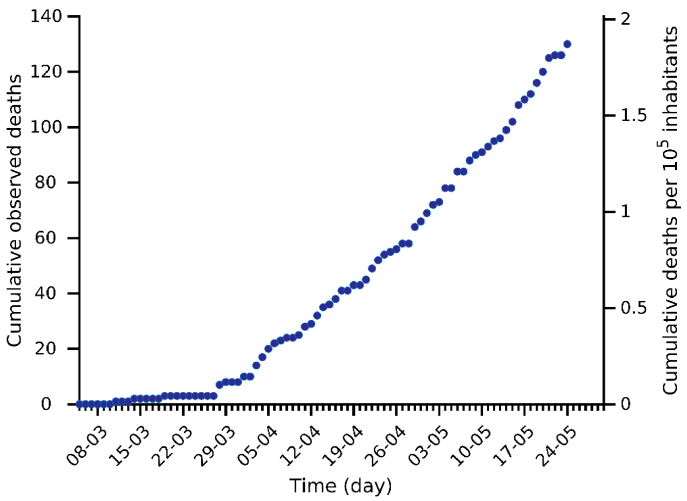


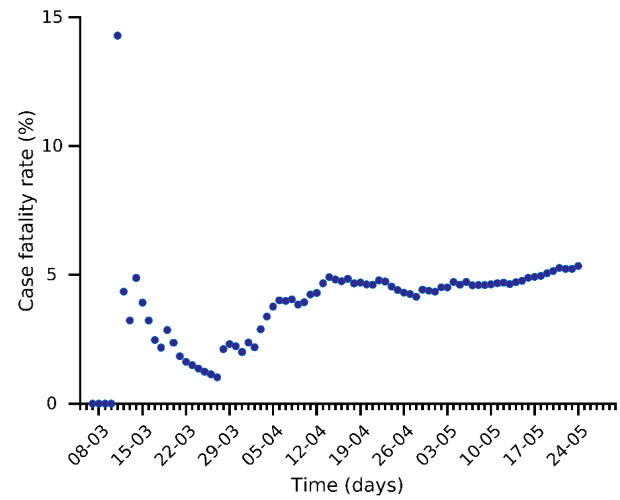




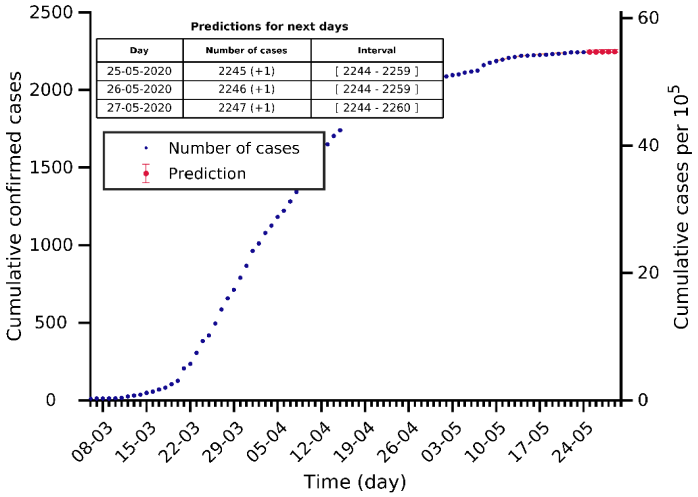


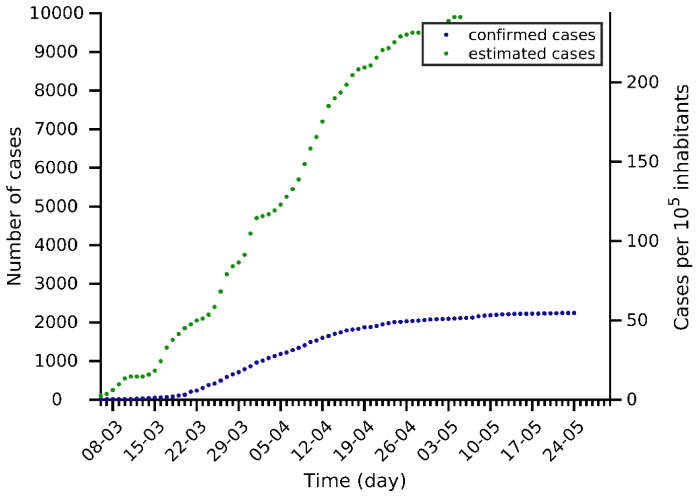


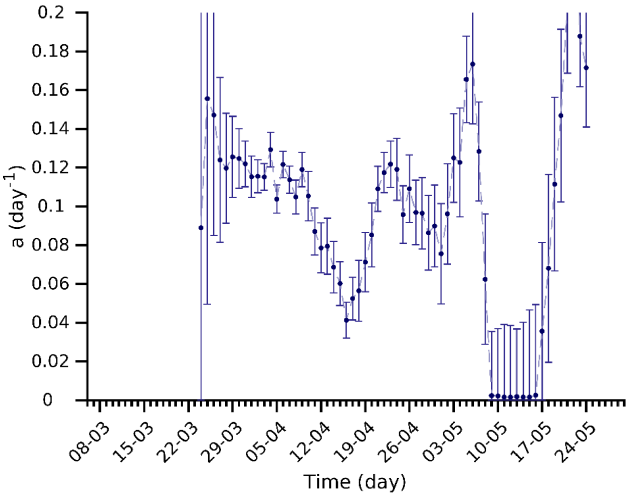


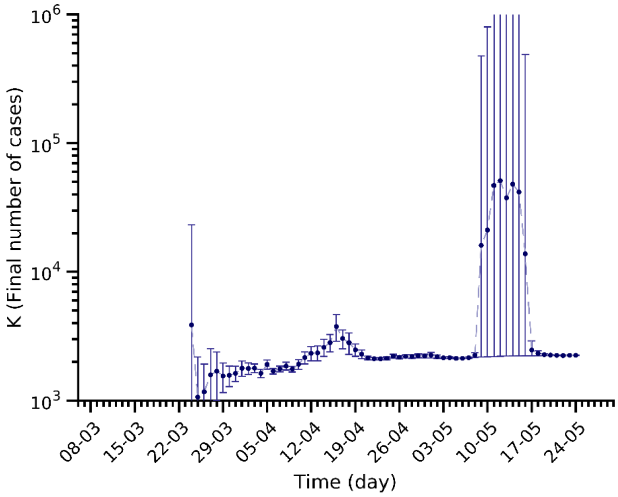


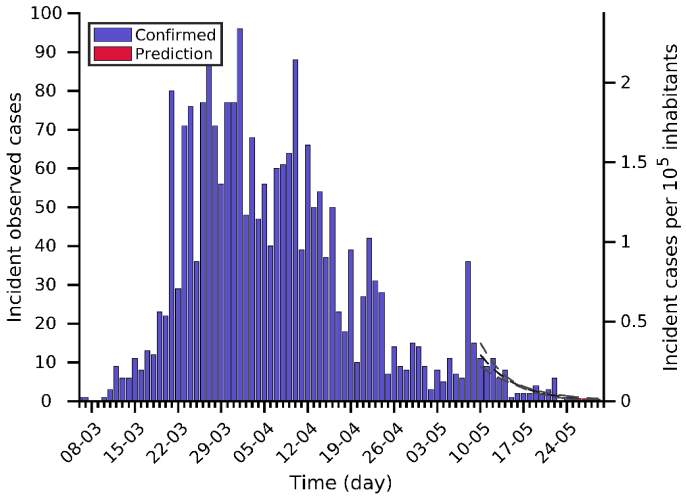
Croatia 24-05-2020. Population: 4.1M. Current cumulated incidence: $55/10^5$

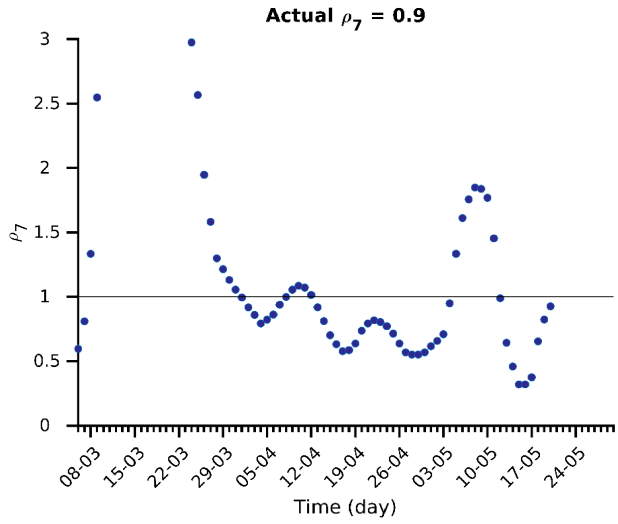


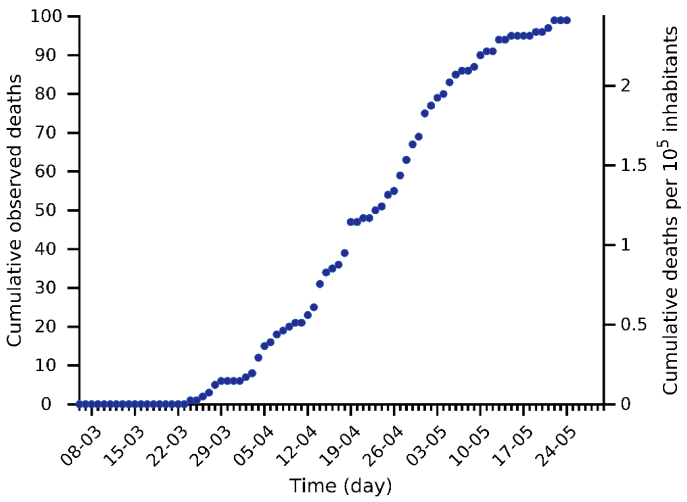


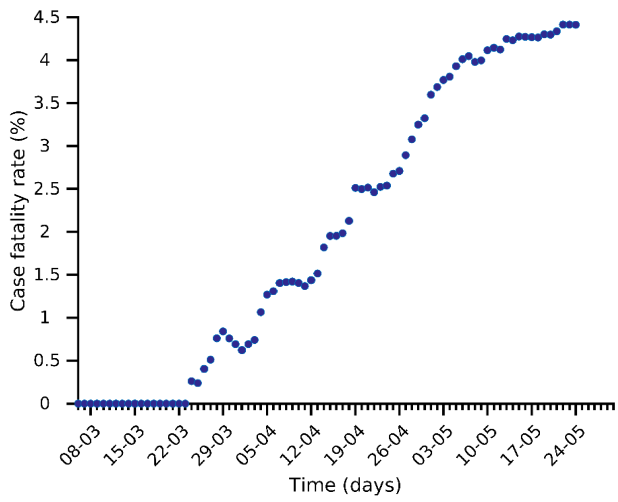




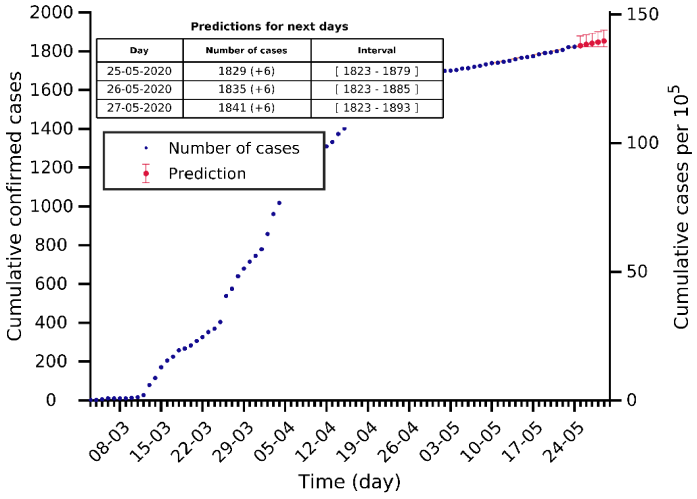


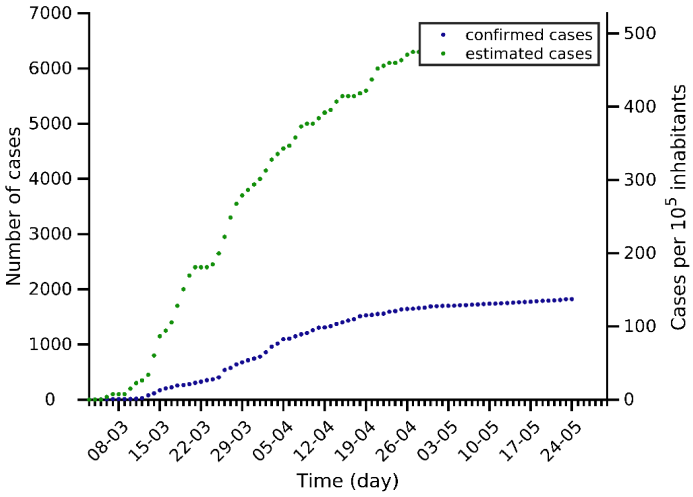


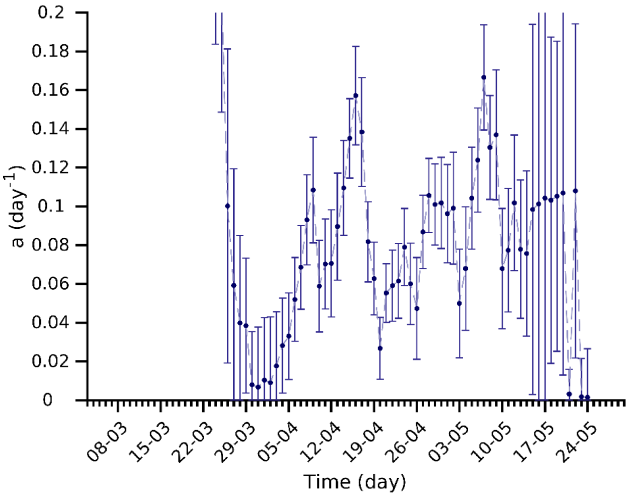


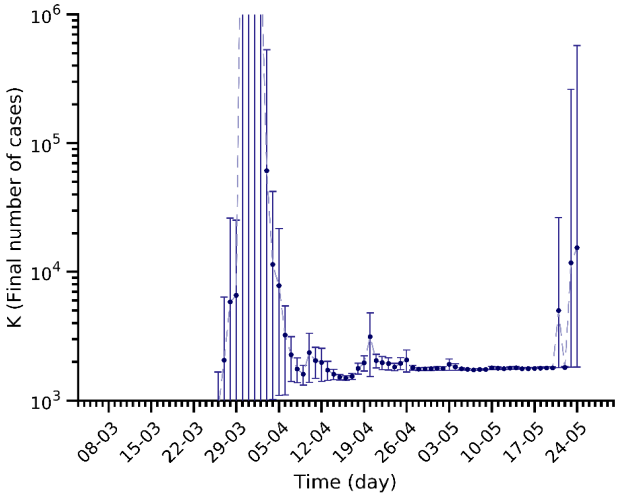


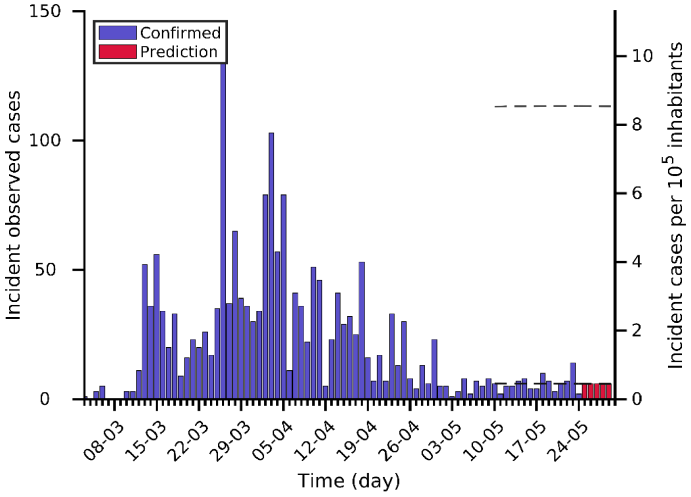
Estonia 24-05-2020. Population: 1.3M. Current cumulated incidence: 137/10⁵

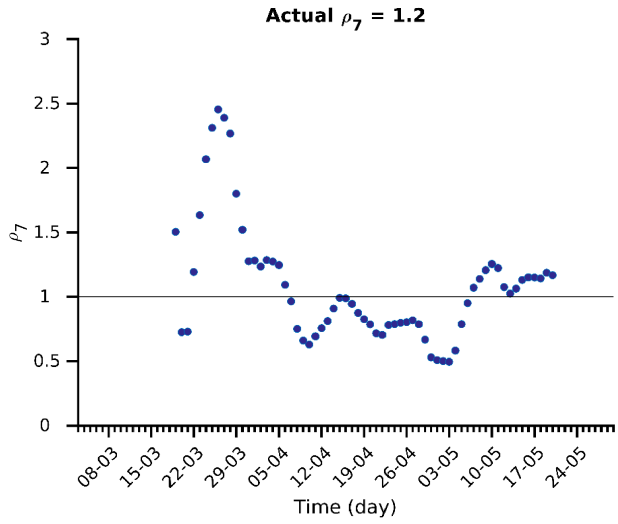


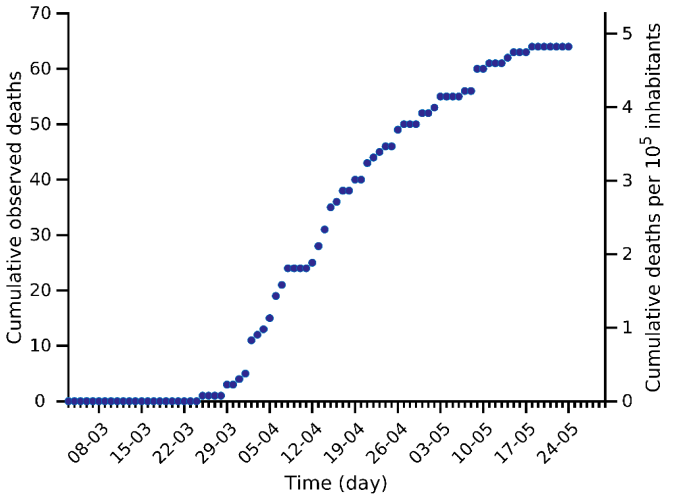


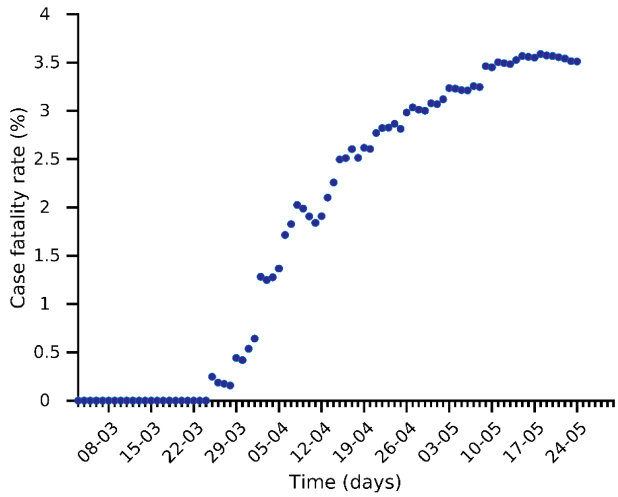




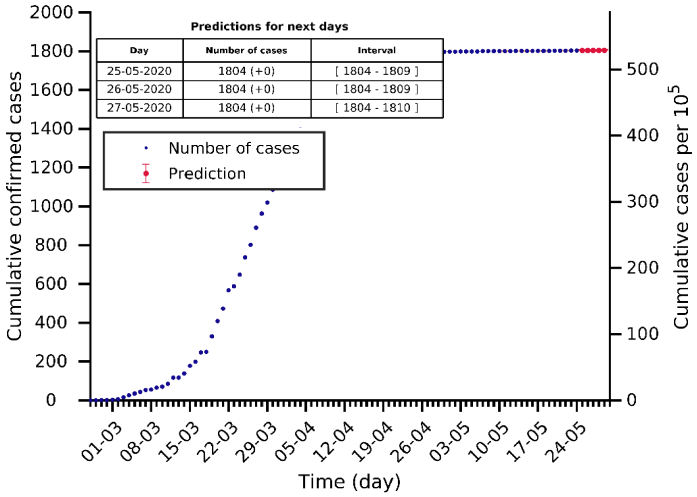


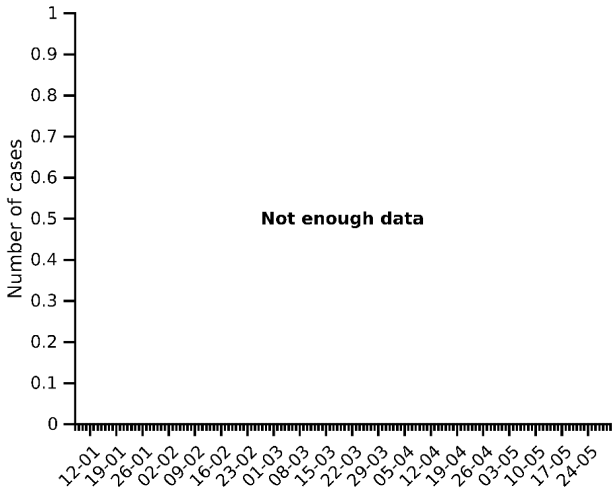


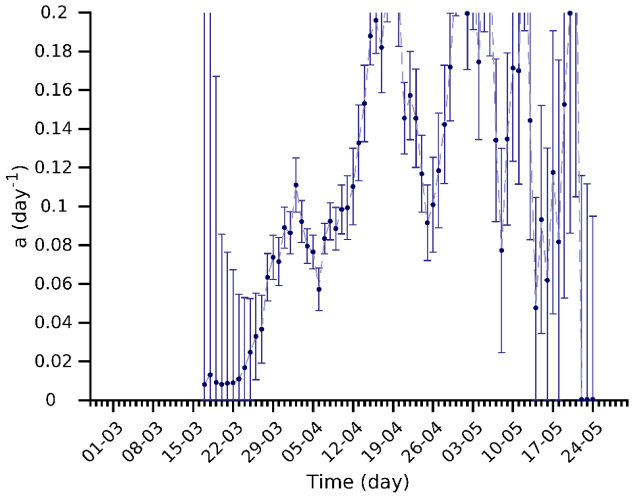


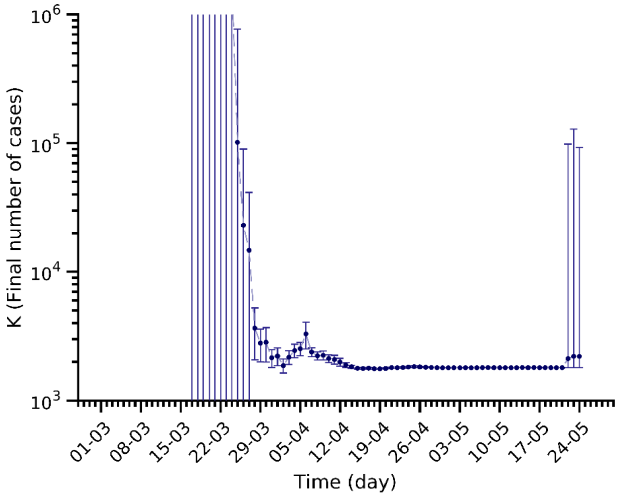


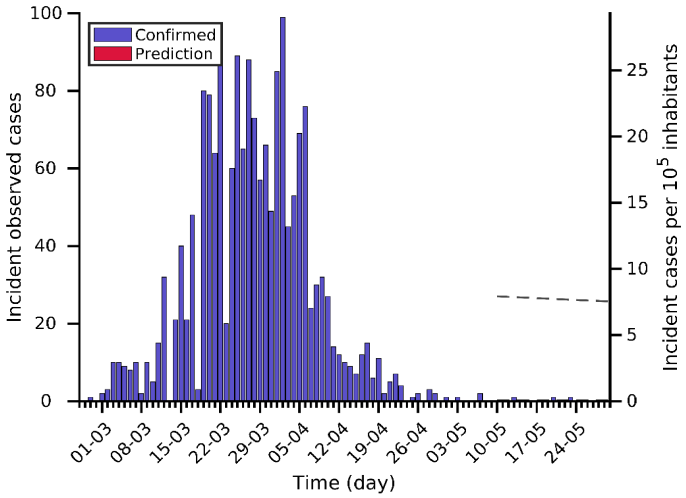
Iceland 24-05-2020. Population: 0.3M. Current cumulated incidence: $529/10^5$

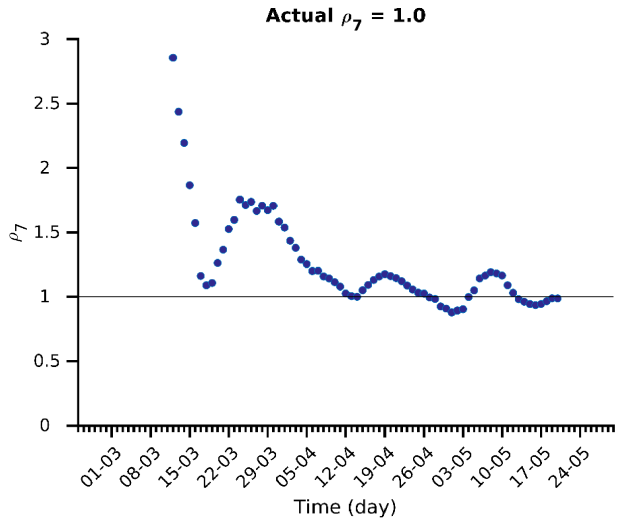


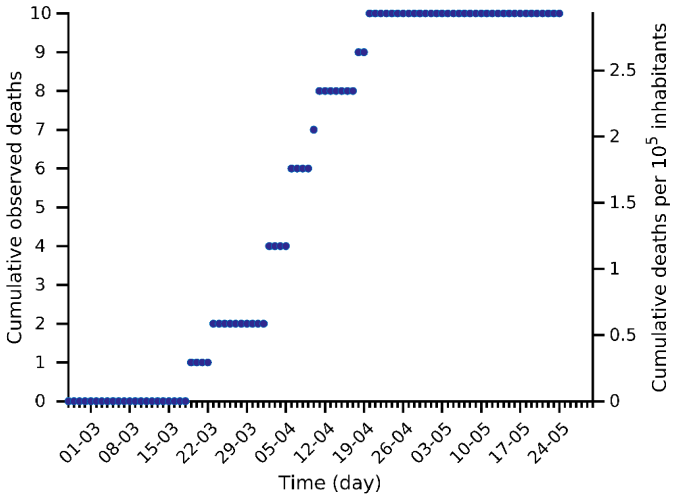


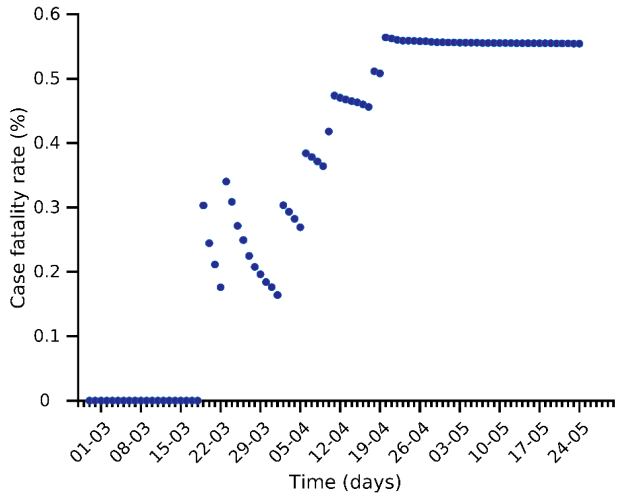




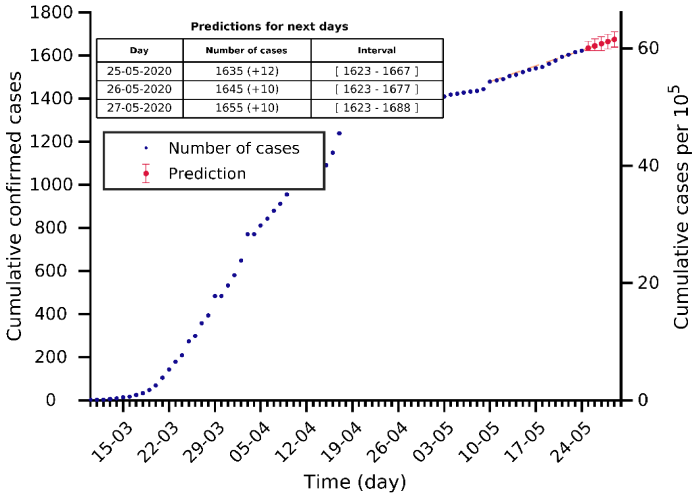


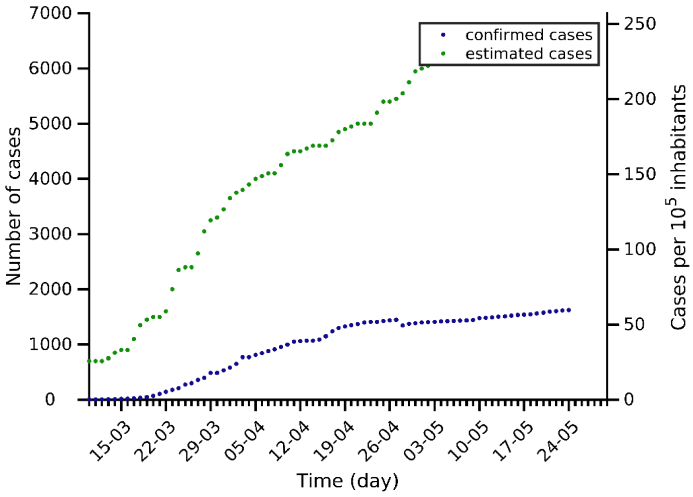


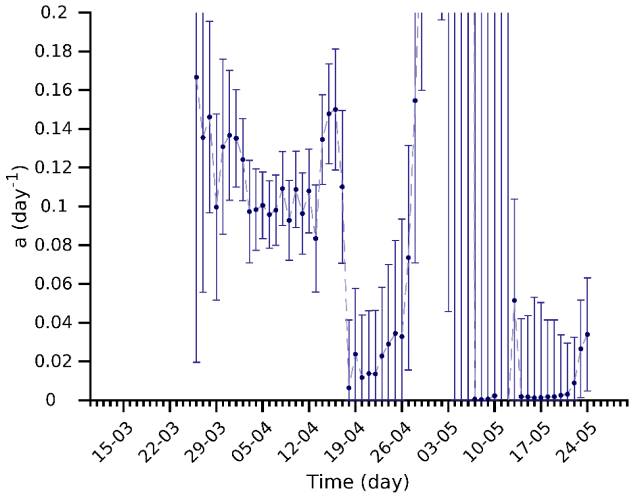


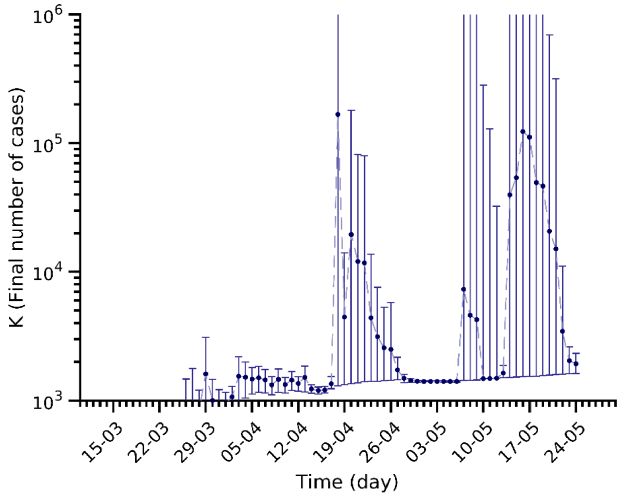


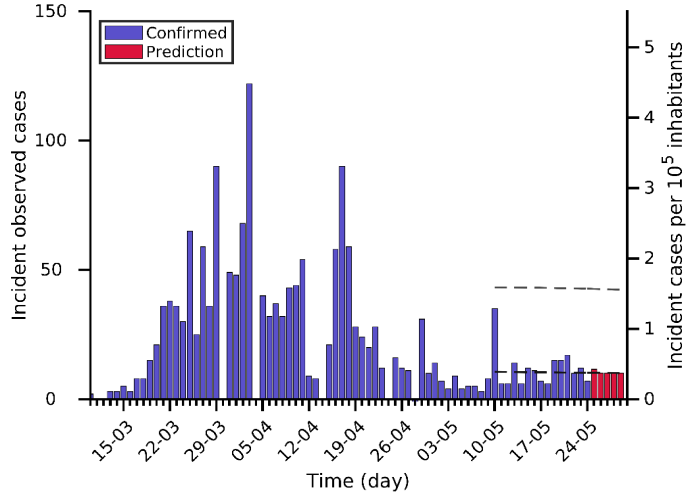
Lithuania 24-05-2020. Population: 2.7M. Current cumulated incidence: $60/10^5$

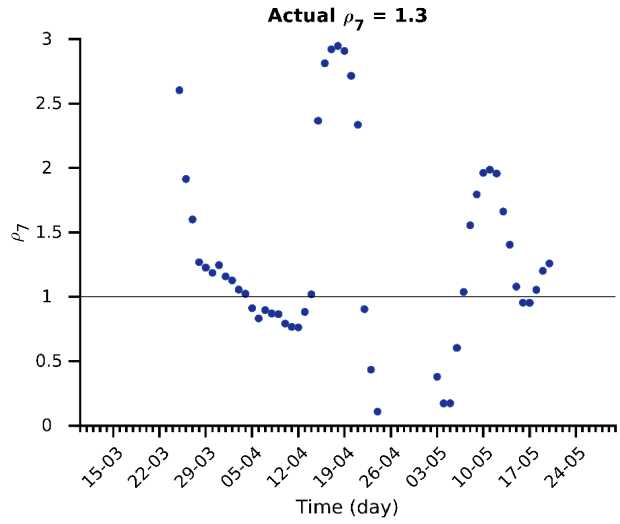


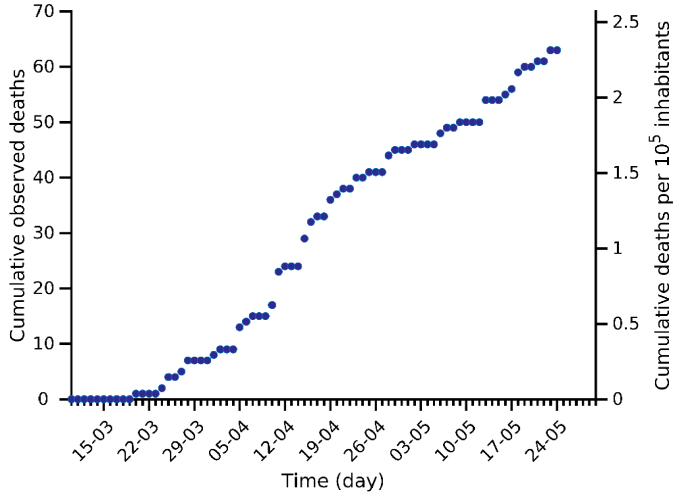


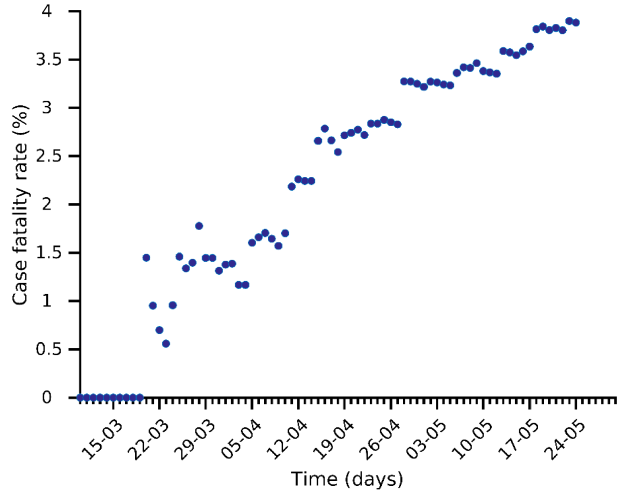




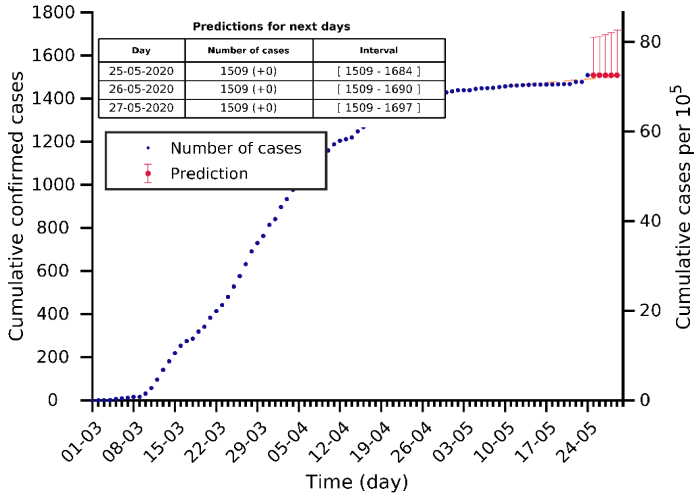


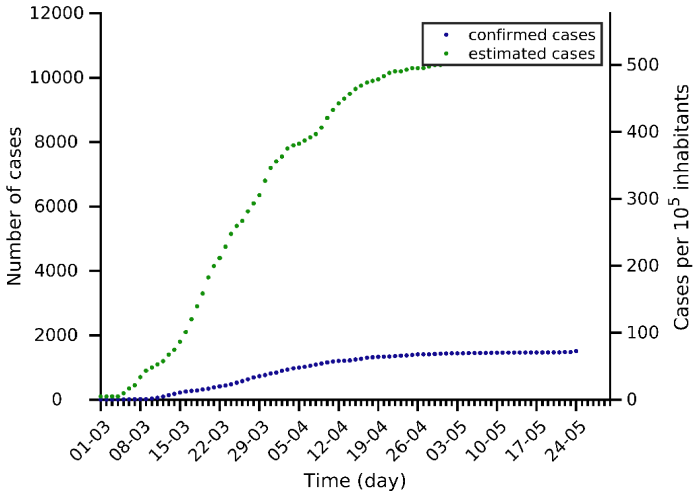


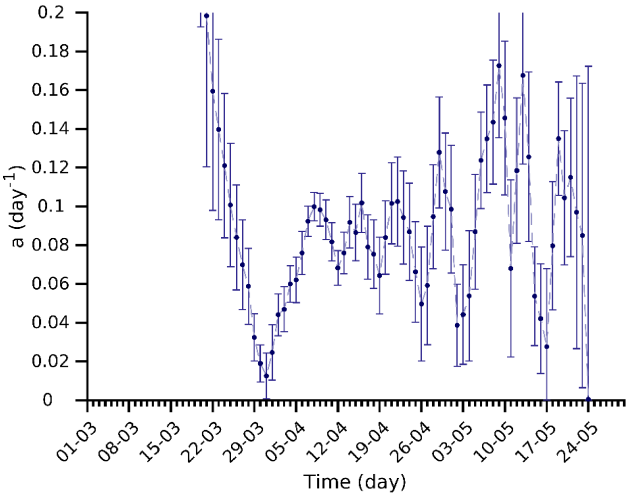


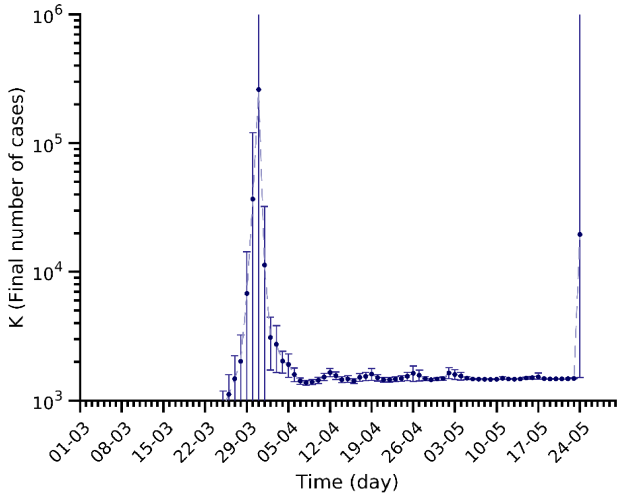


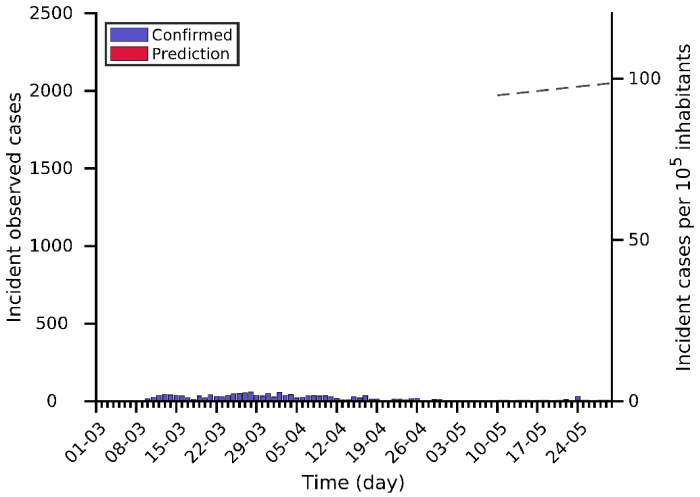
Slovenia 24-05-2020. Population: 2.1M. Current cumulated incidence: 73/10^5

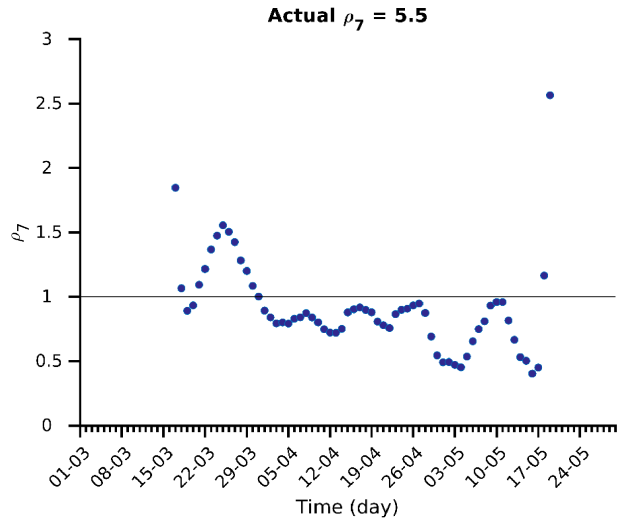


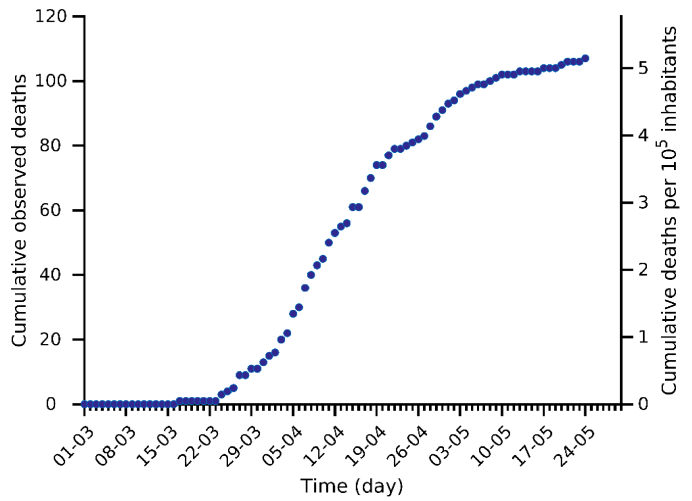


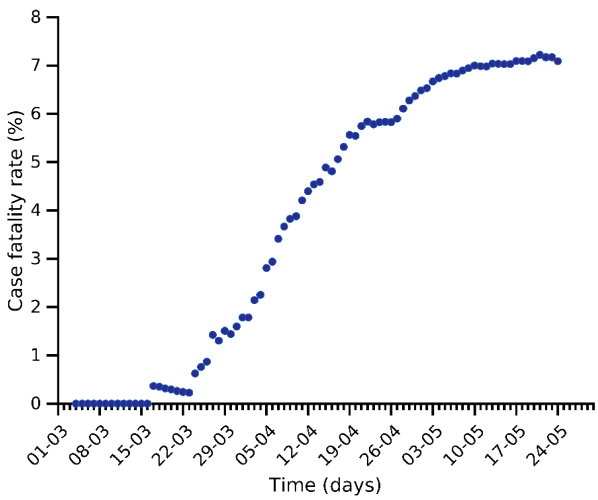




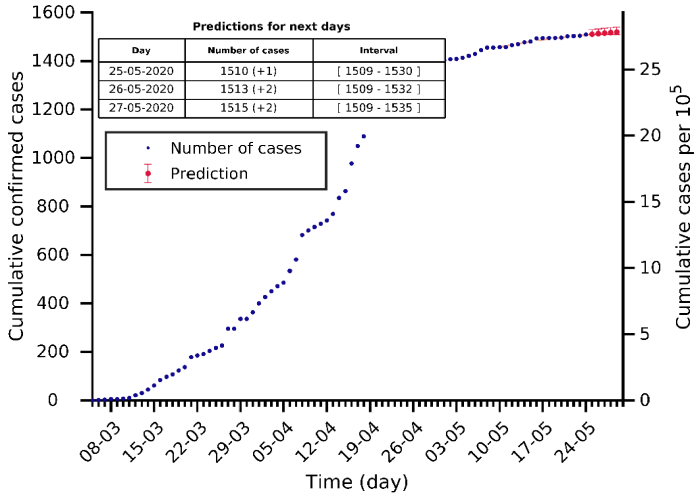


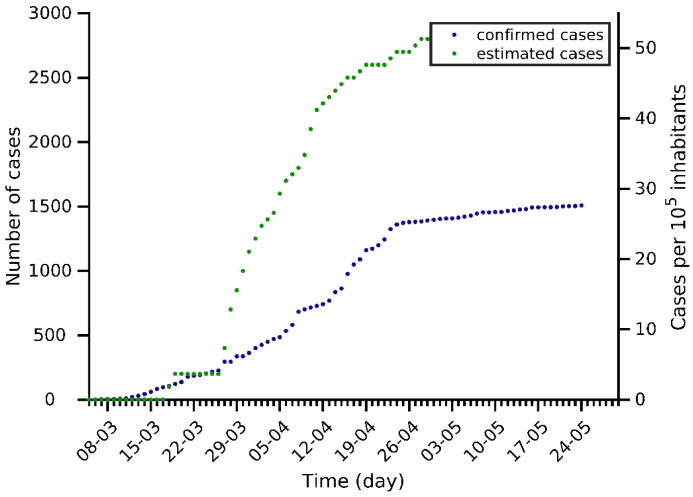


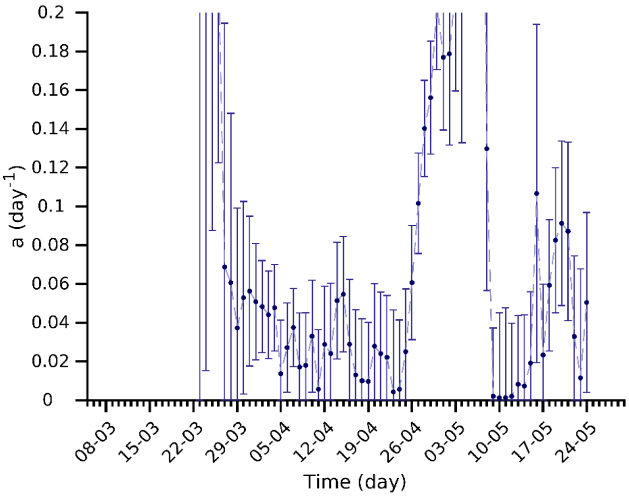


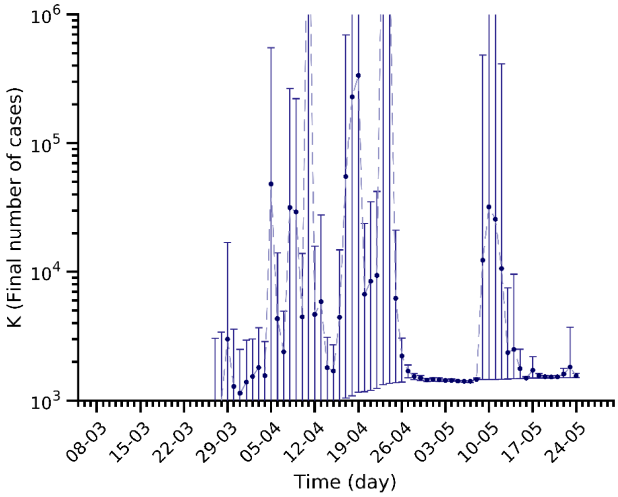


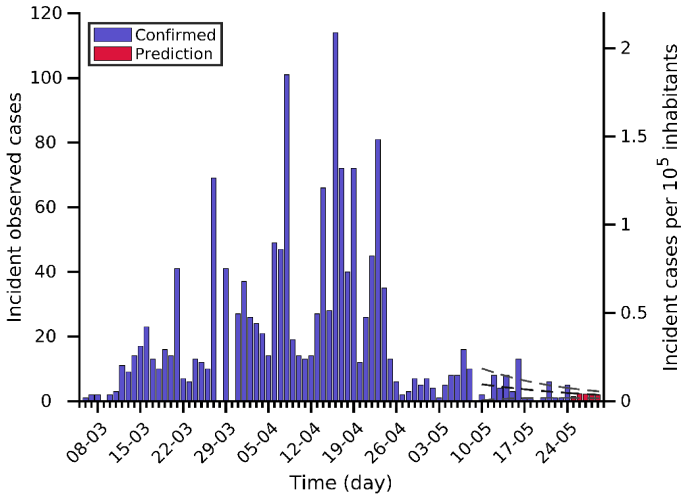
Slovakia 24-05-2020. Population: 5.5M. Current cumulated incidence: 28/10^5

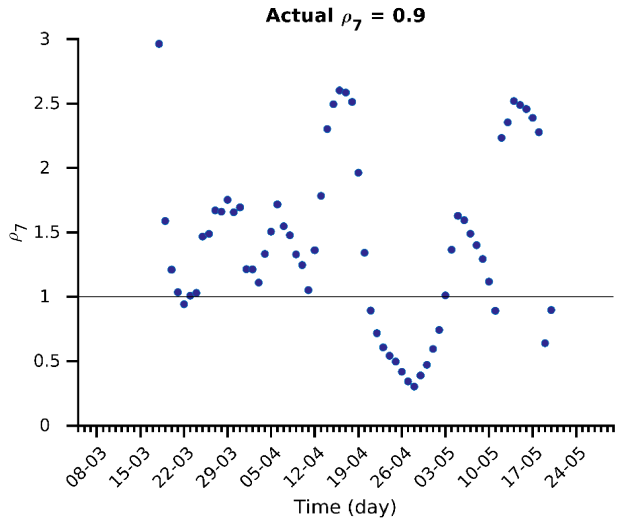


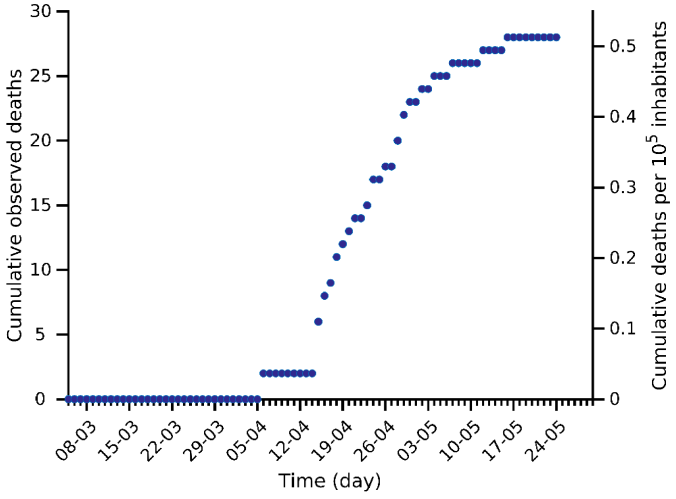


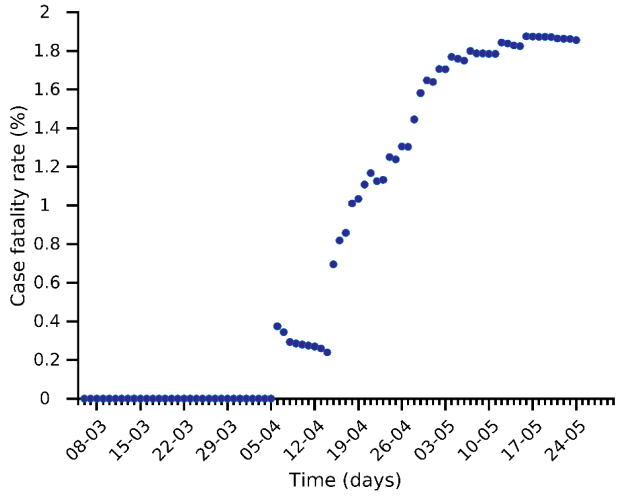




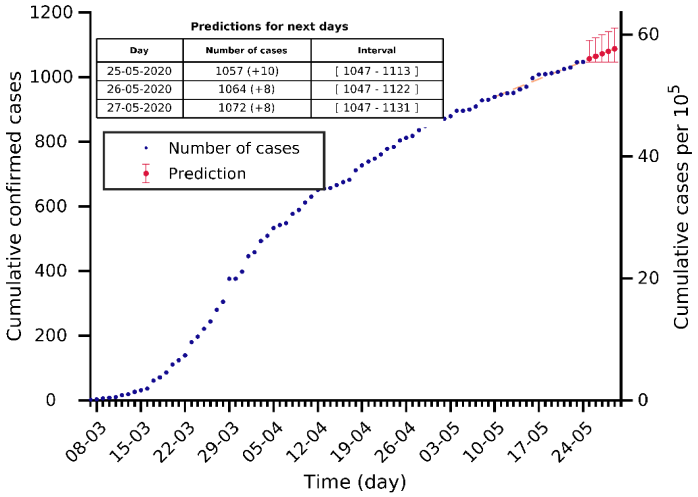


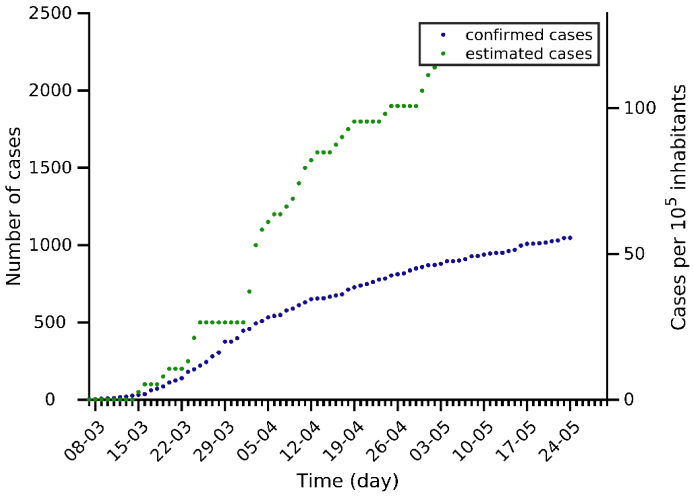


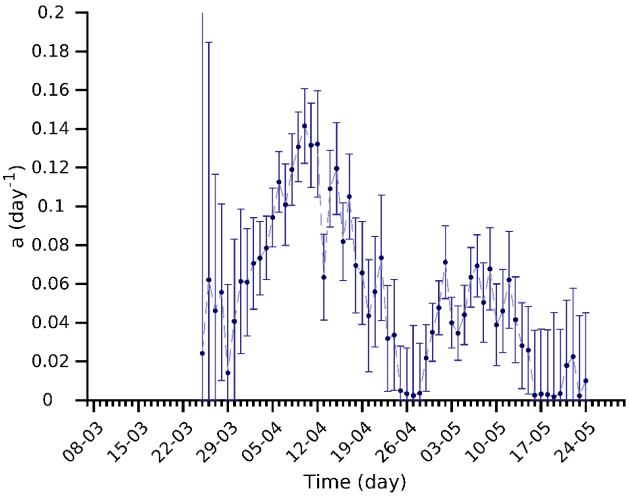


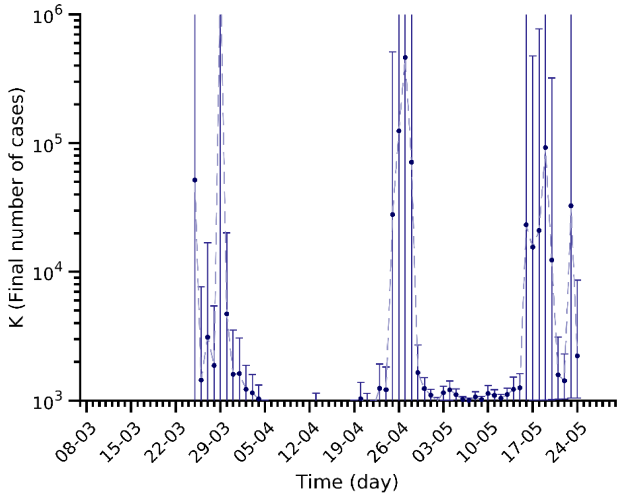


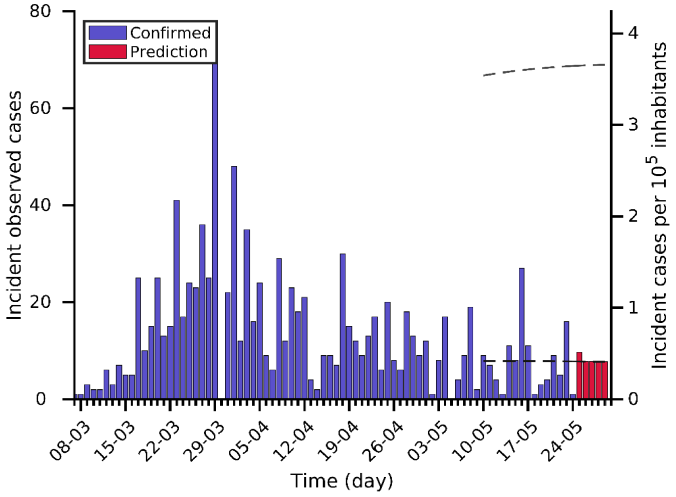
Latvia 24-05-2020. Population: 1.9M. Current cumulated incidence: $56/10^5$

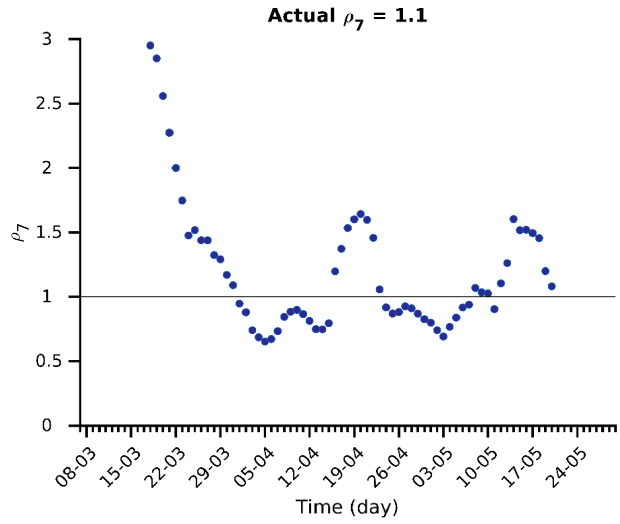


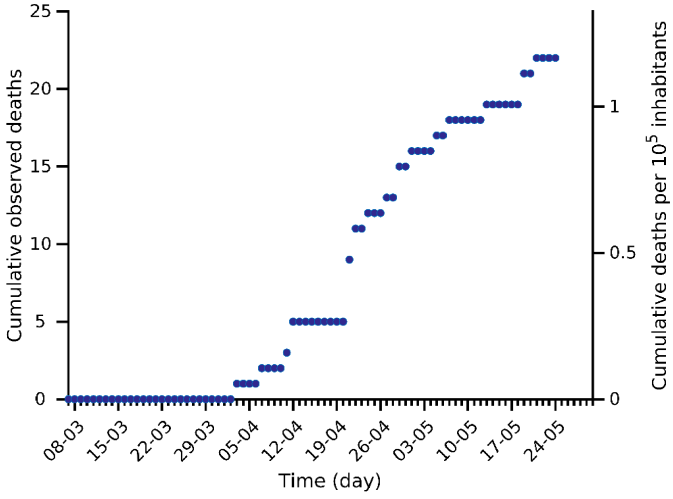


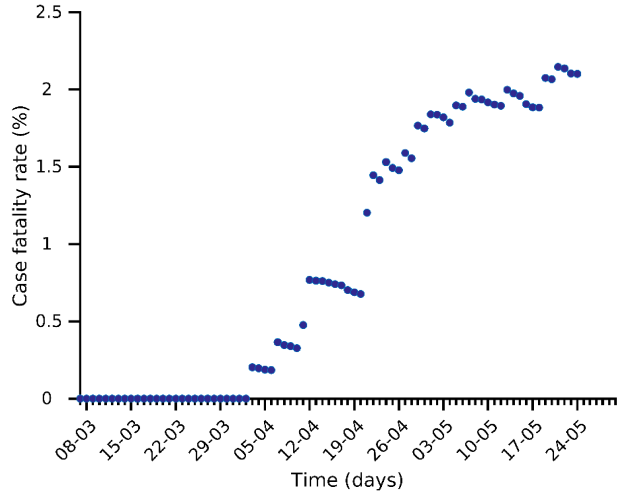




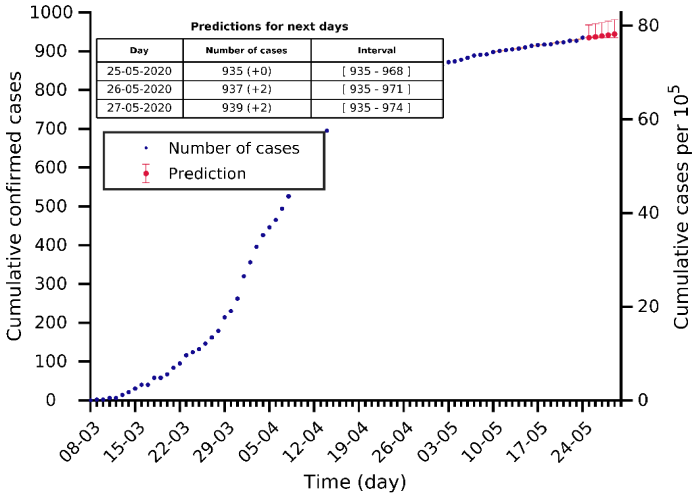


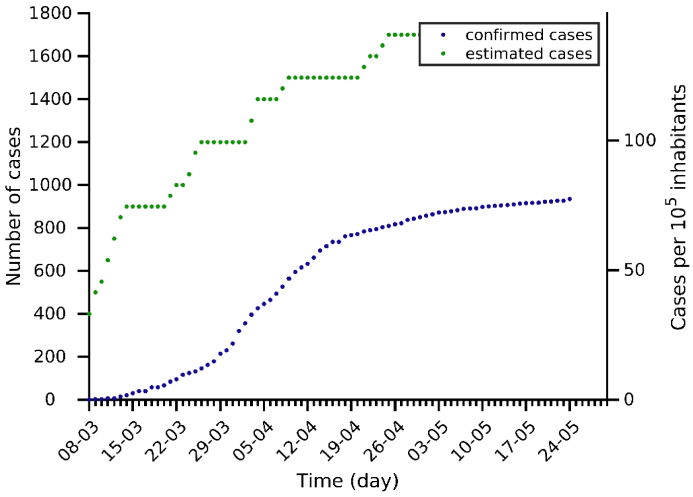


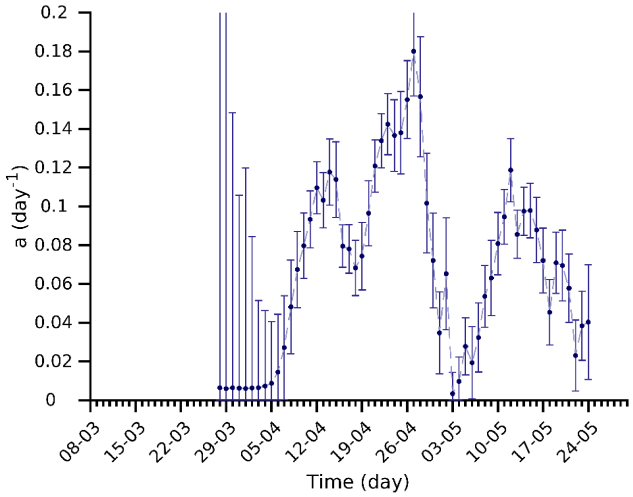


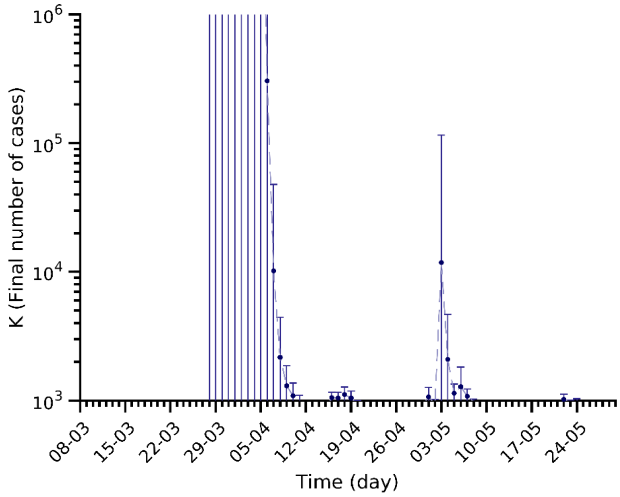


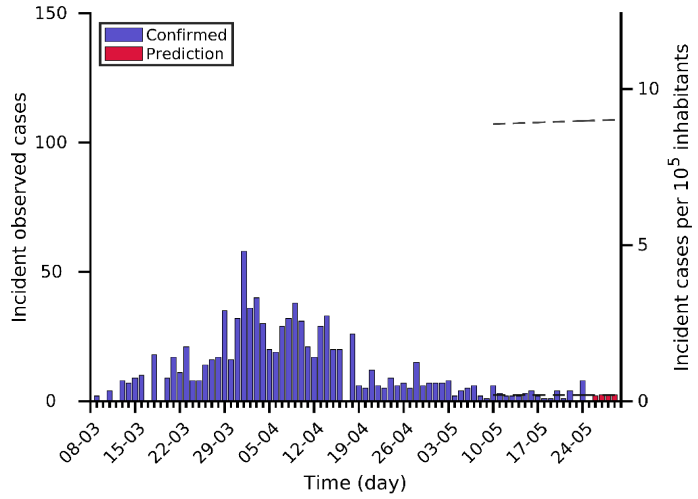
Cyprus 24-05-2020. Population: 1.2M. Current cumulated incidence: $77/10^5$

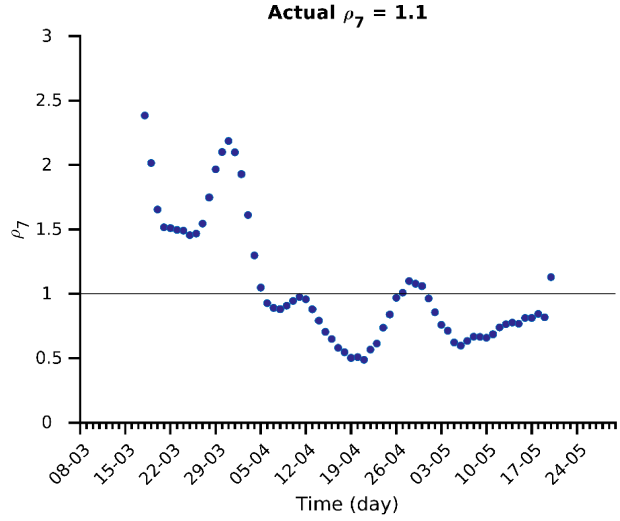


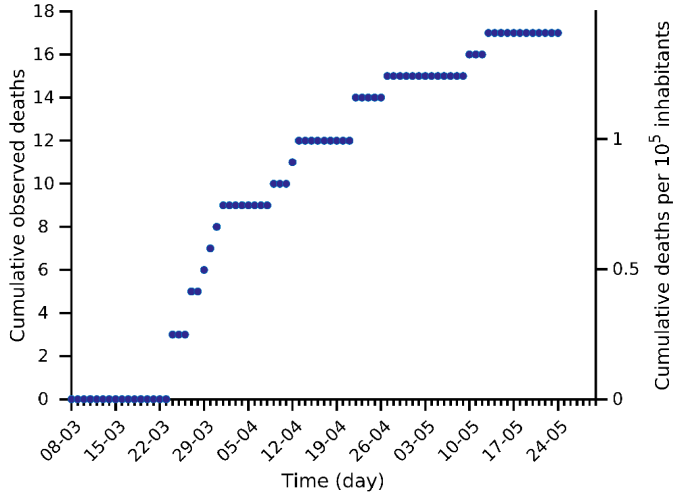


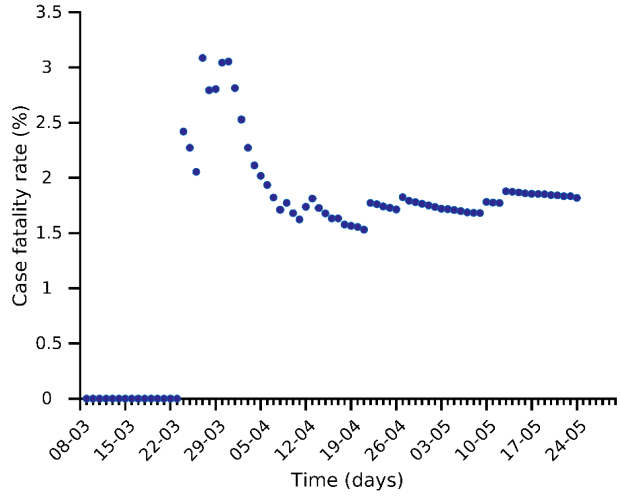




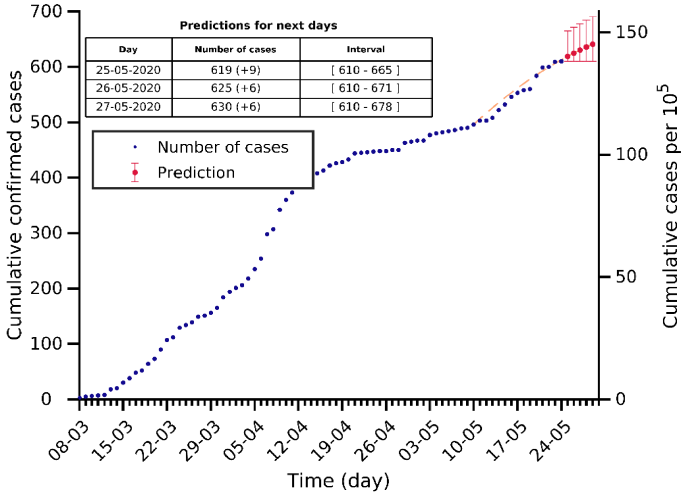


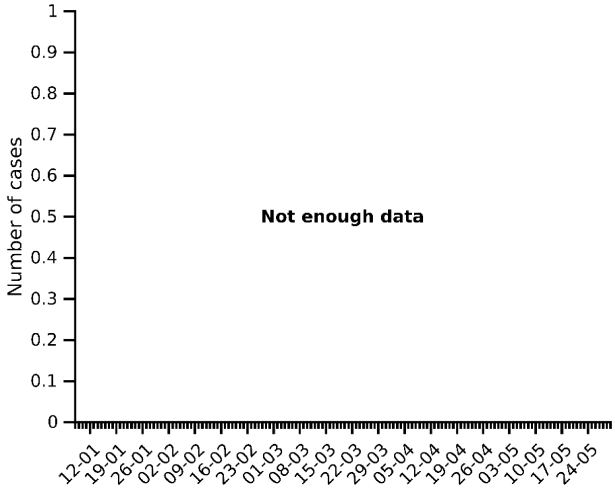


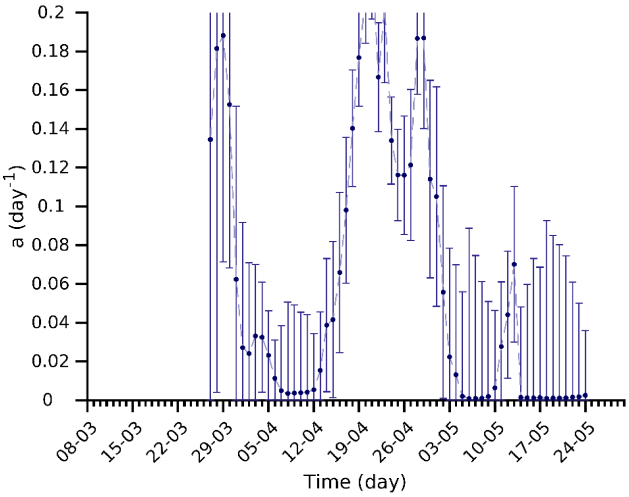


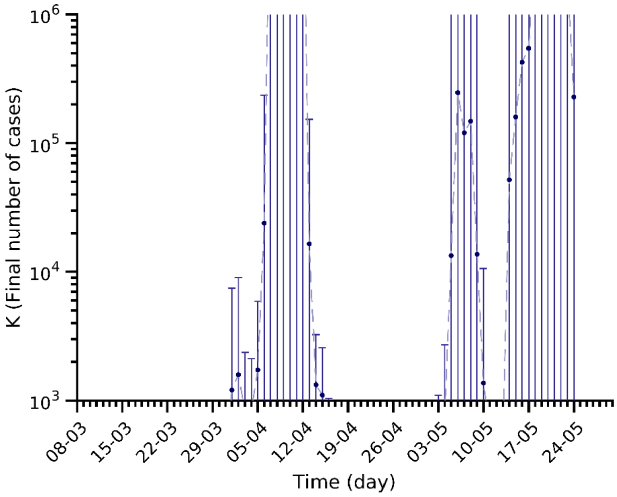


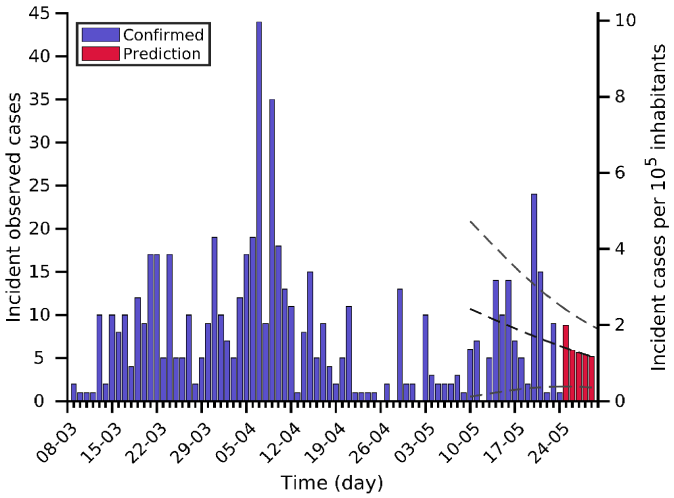
Malta 24-05-2020. Population: 0.4M. Current cumulated incidence: 138/10^5

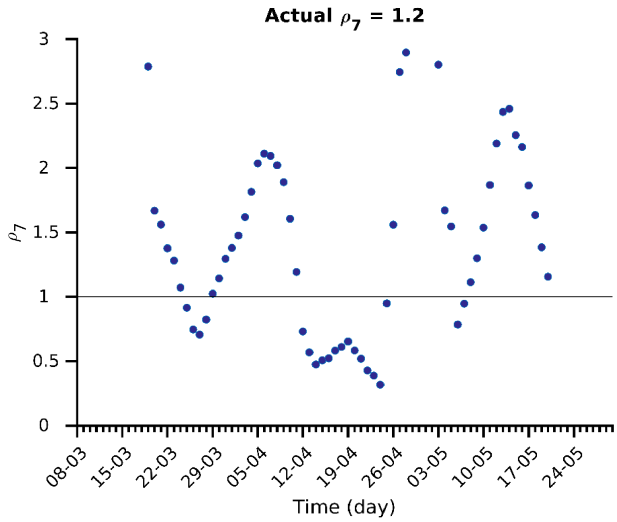


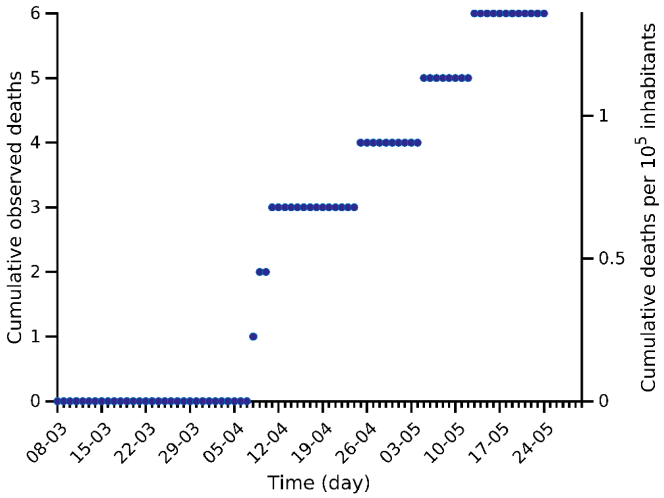


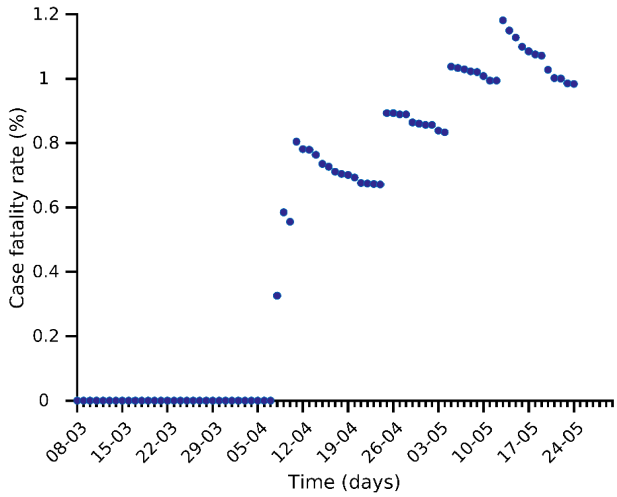








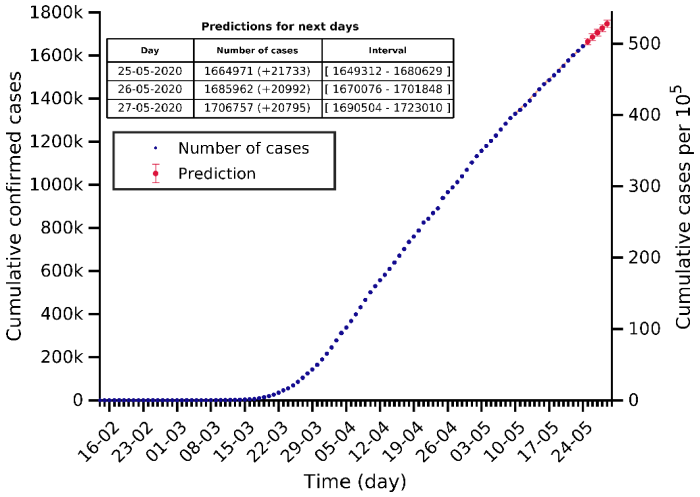


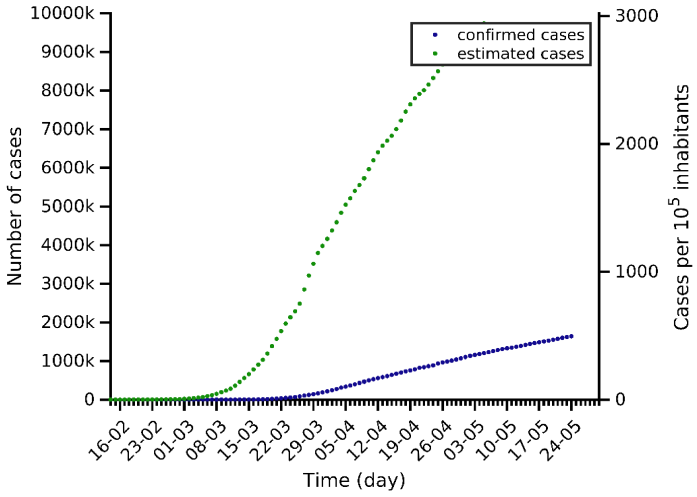


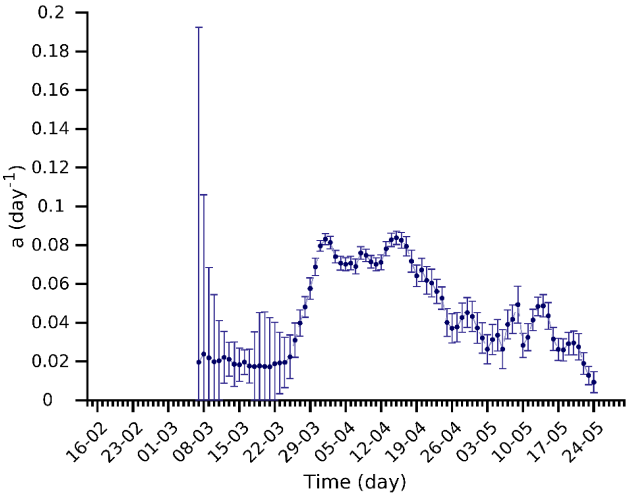
(2) Analysis and prediction of COVID-19 for other countries

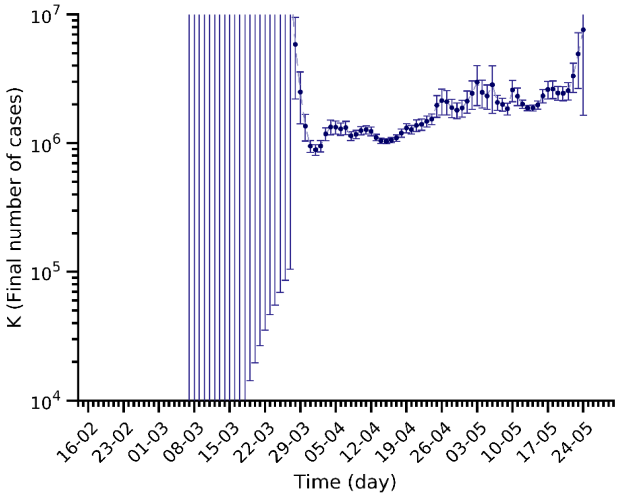
Data obtained from https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

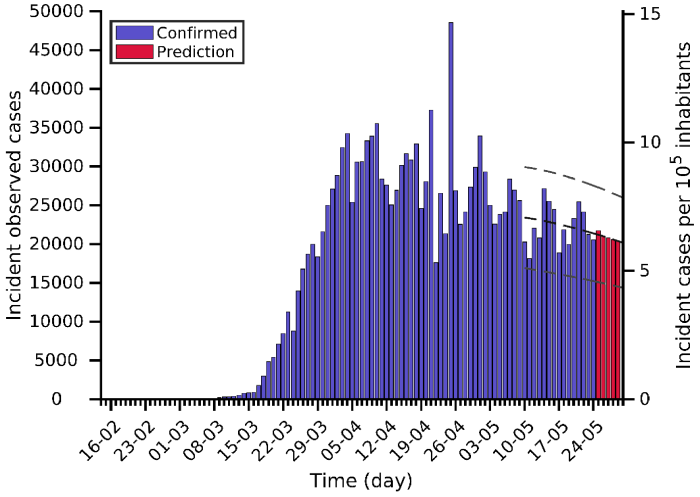
USA 24-05-2020. Population: 331.0M. Current cumulated incidence: 496/10⁵

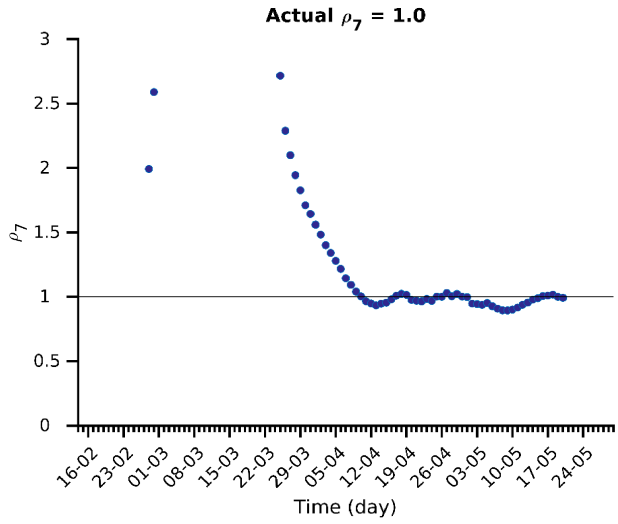


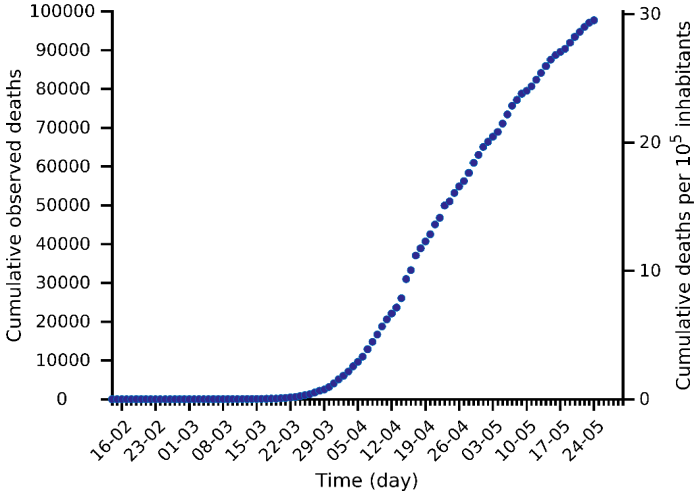


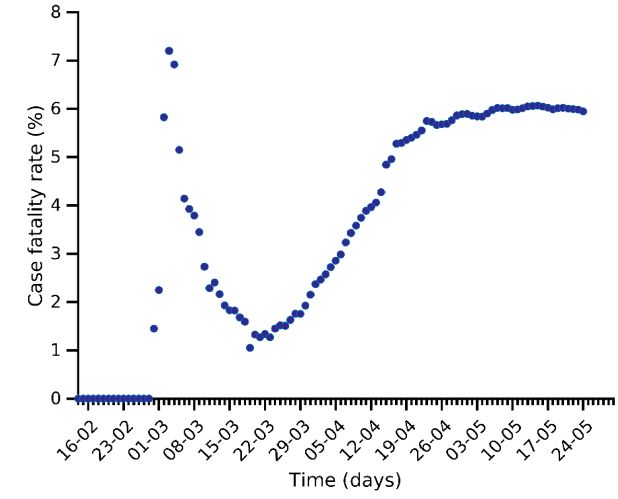




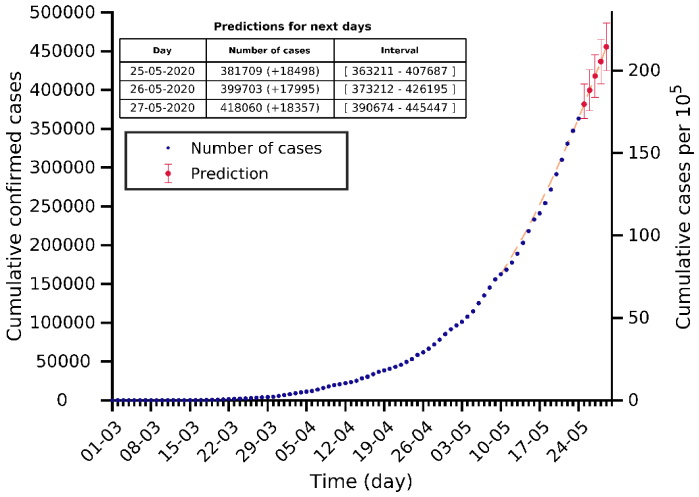


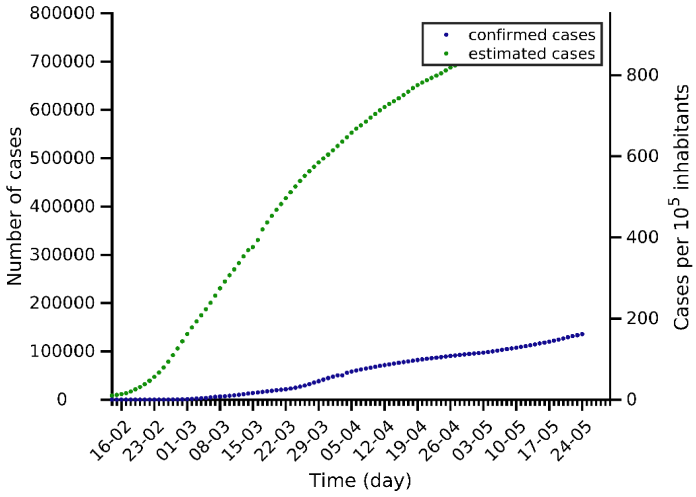


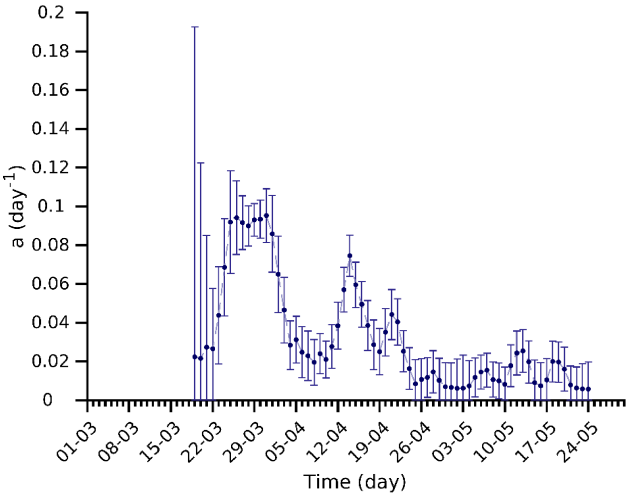


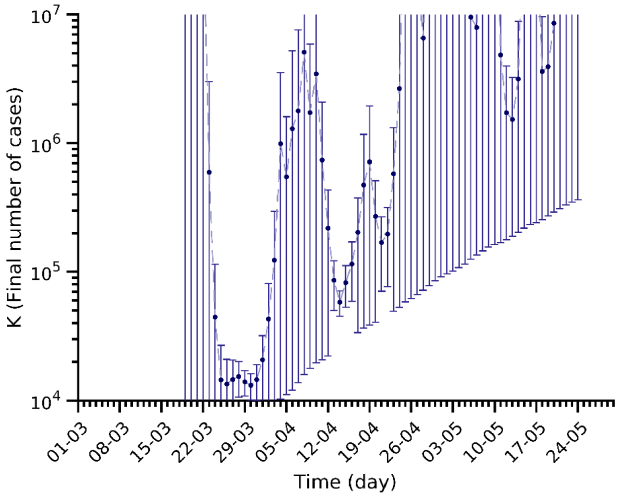


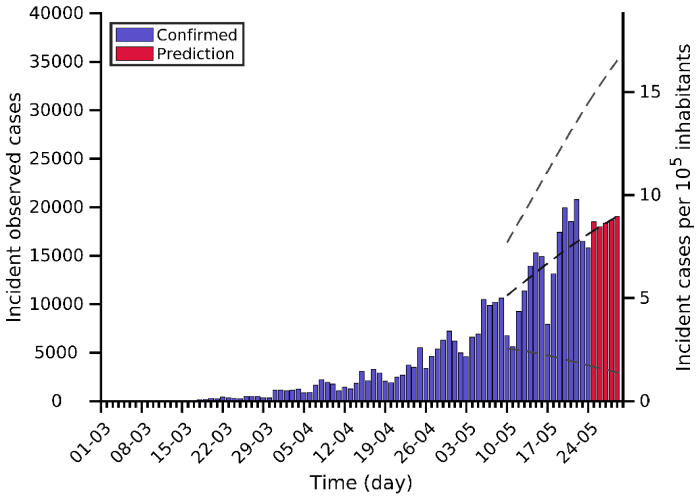
Brazil 24-05-2020. Population: 212.6M. Current cumulated incidence: 171/10^5

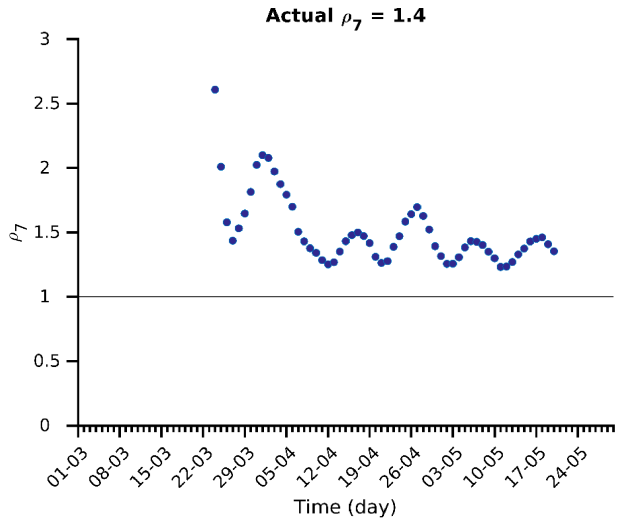


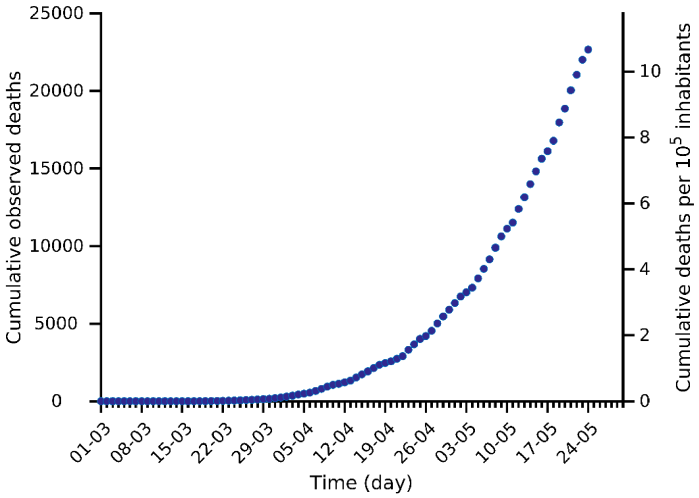


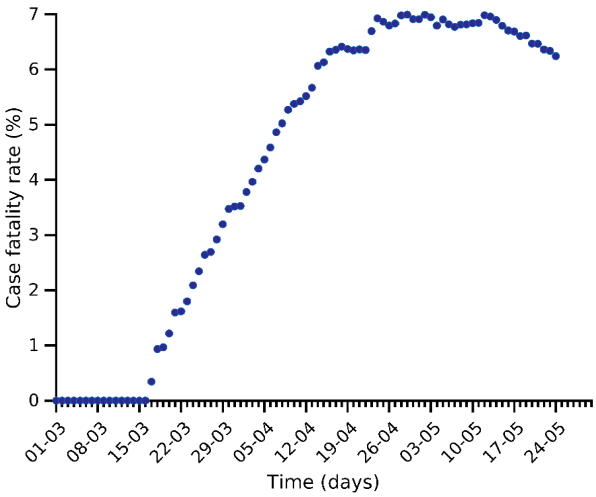




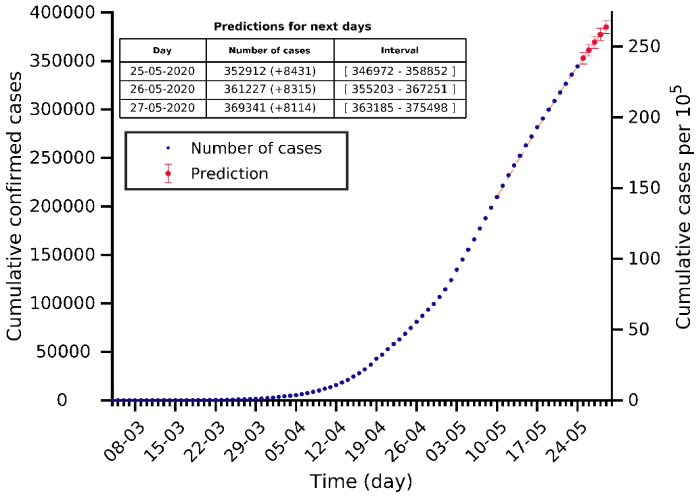


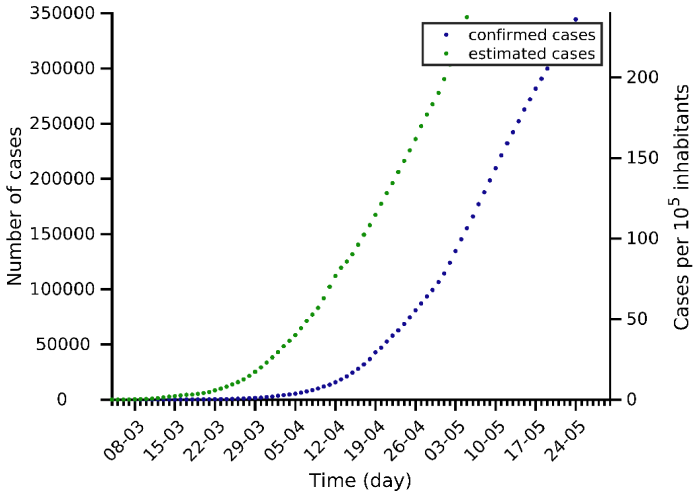


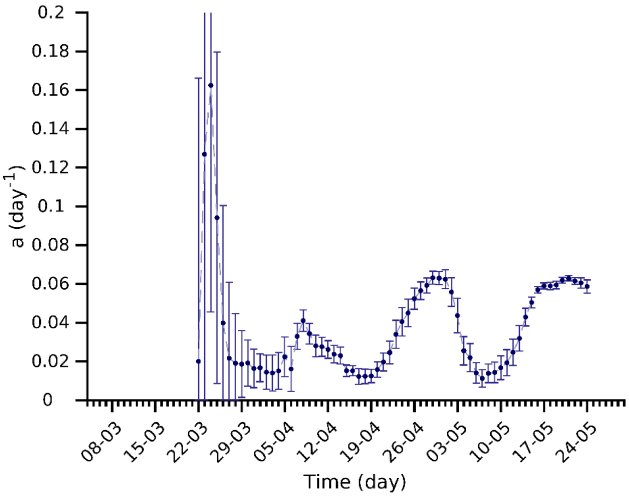


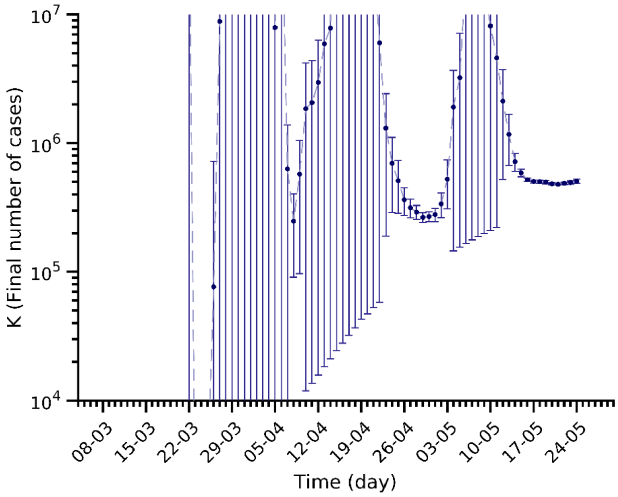


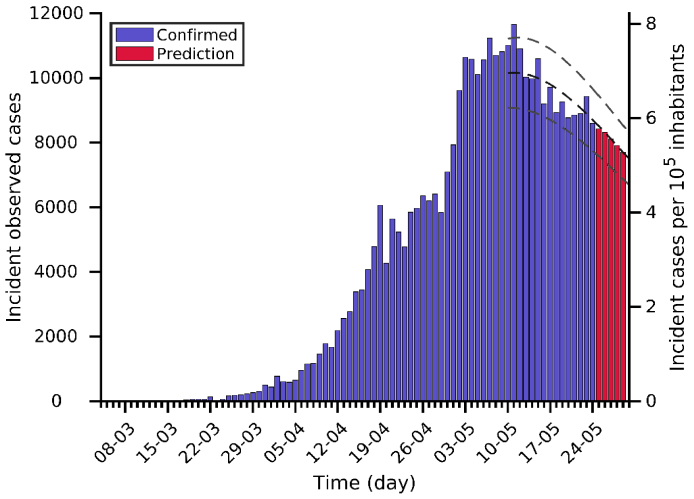
Russia 24-05-2020. Population: 145.9M. Current cumulated incidence: $236/10^5$

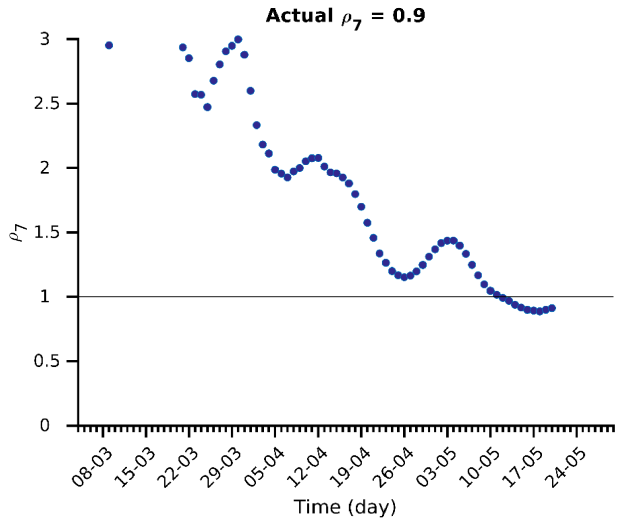


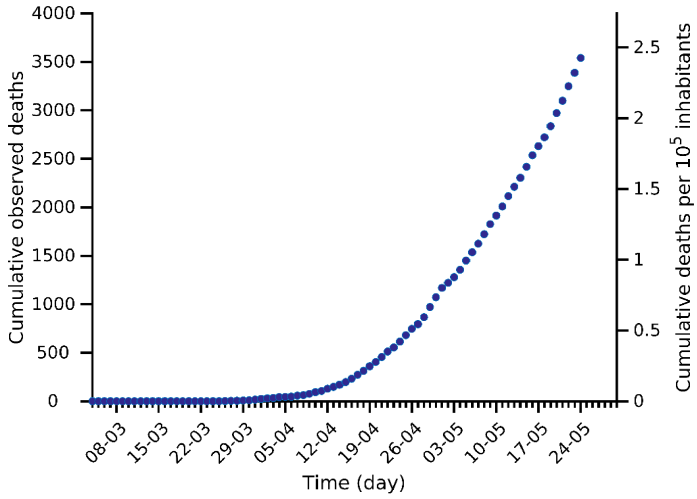


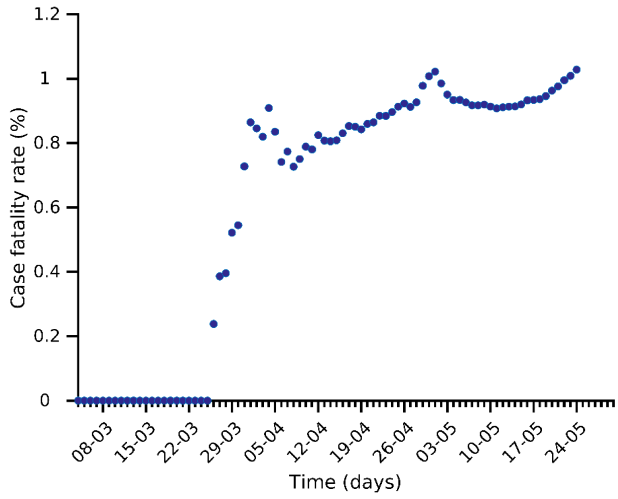




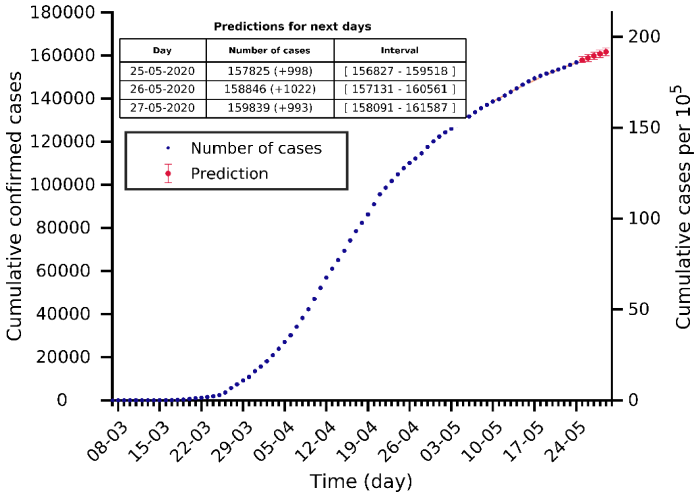


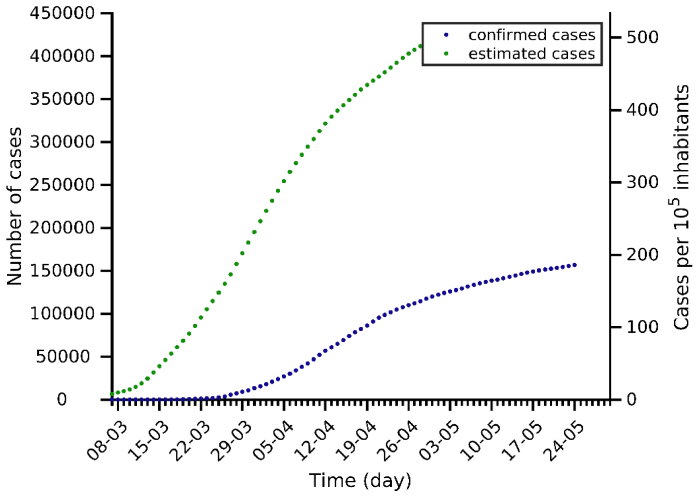


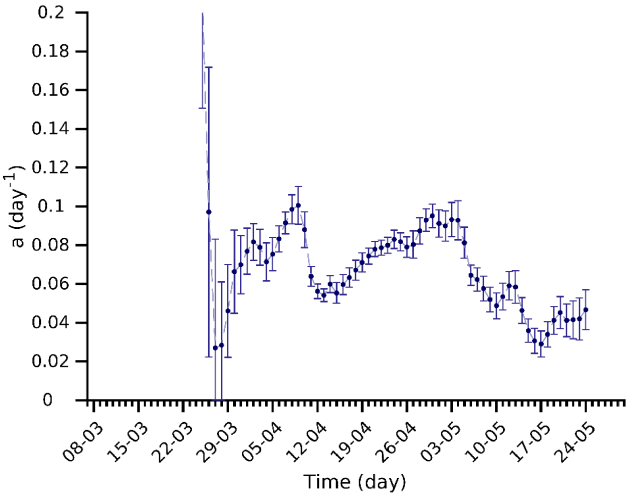


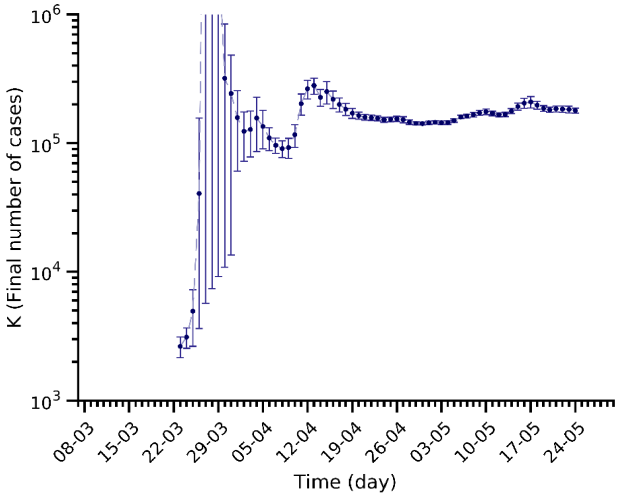


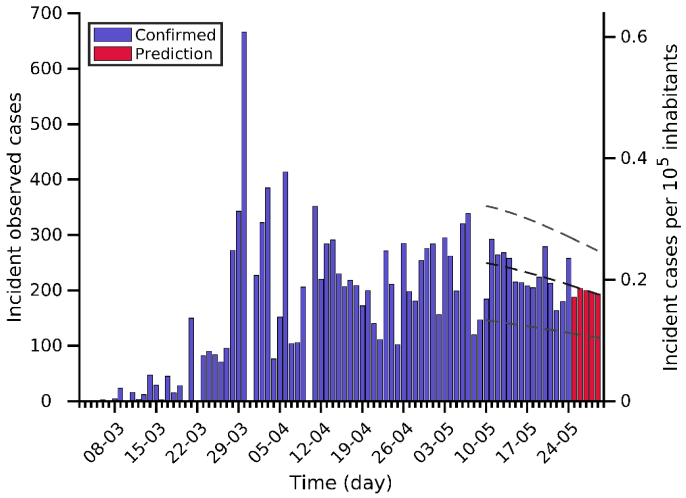
Turkey 24-05-2020. Population: 84.3M. Current cumulated incidence: 186/10^5

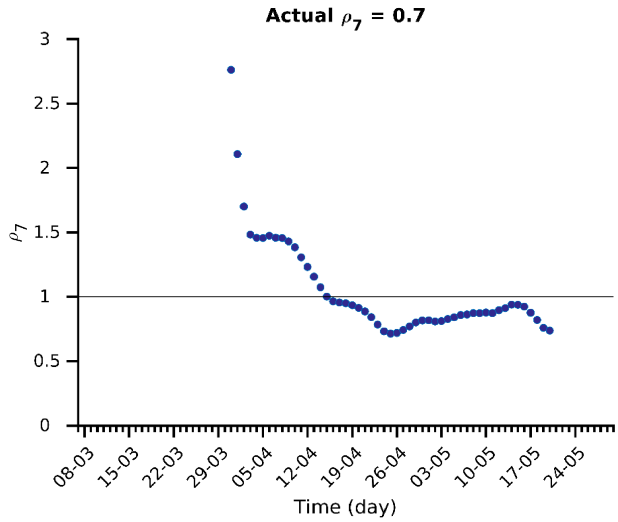


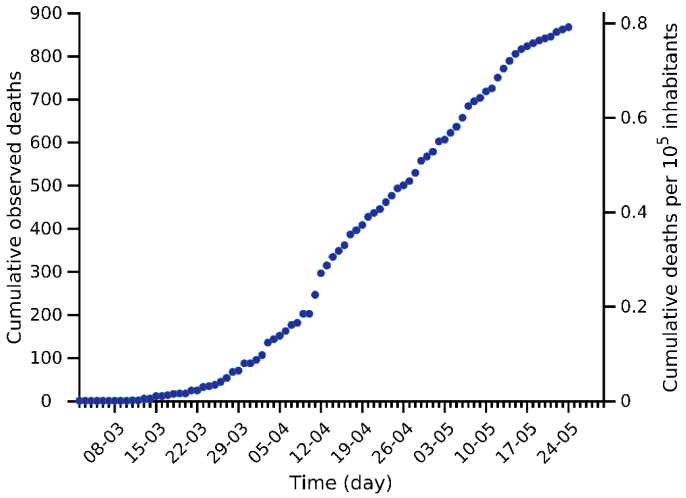


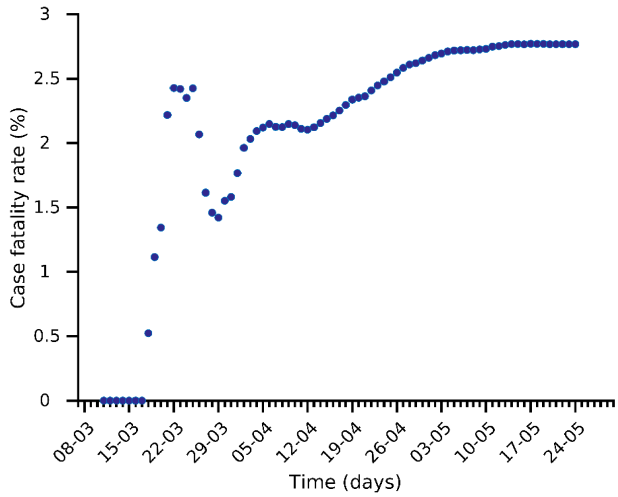




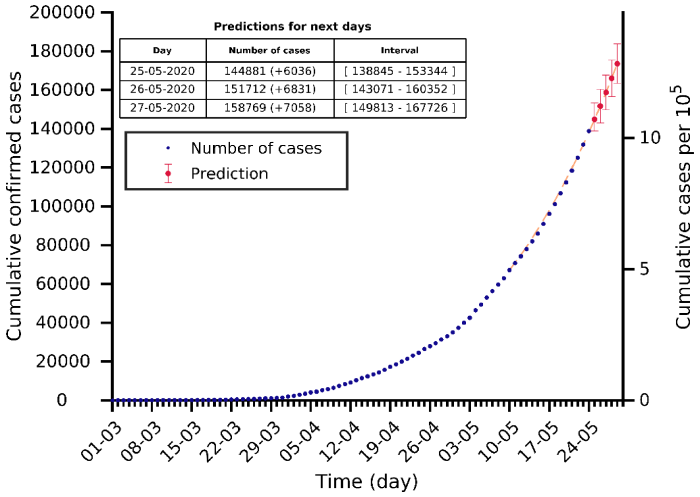


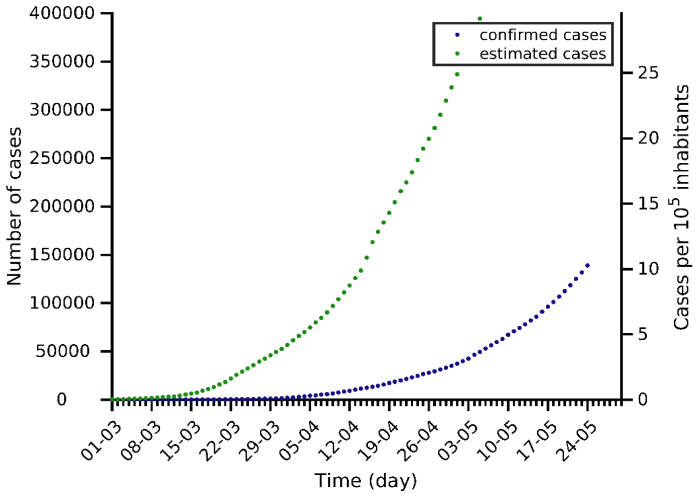


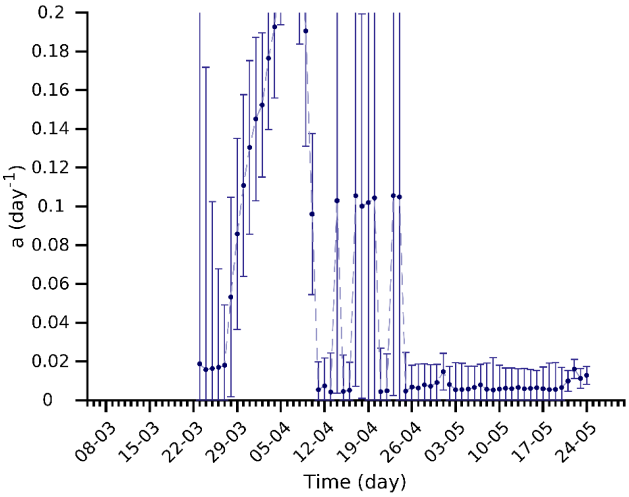


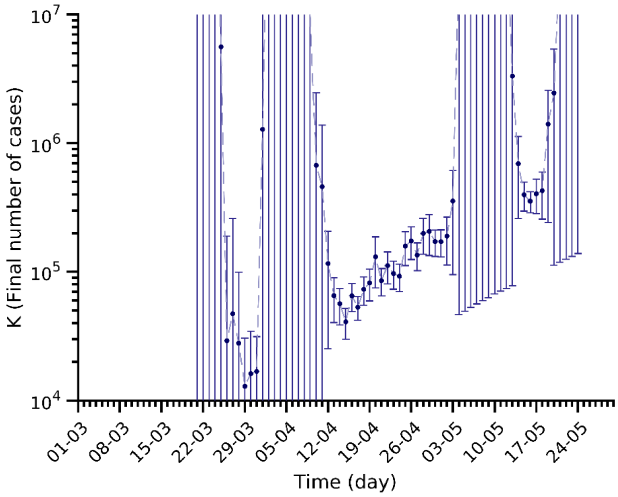


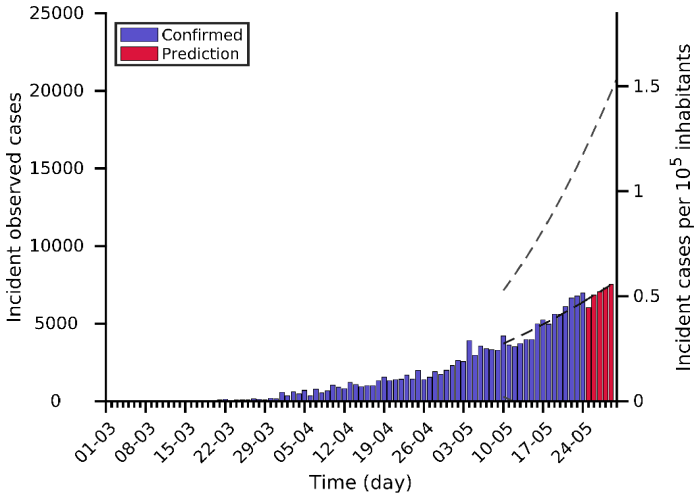
India 24-05-2020. Population: 1353.0M. Current cumulated incidence: $10/10^5$

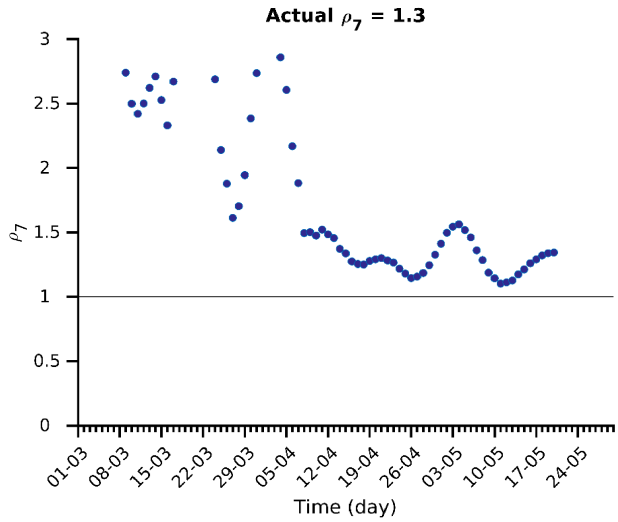


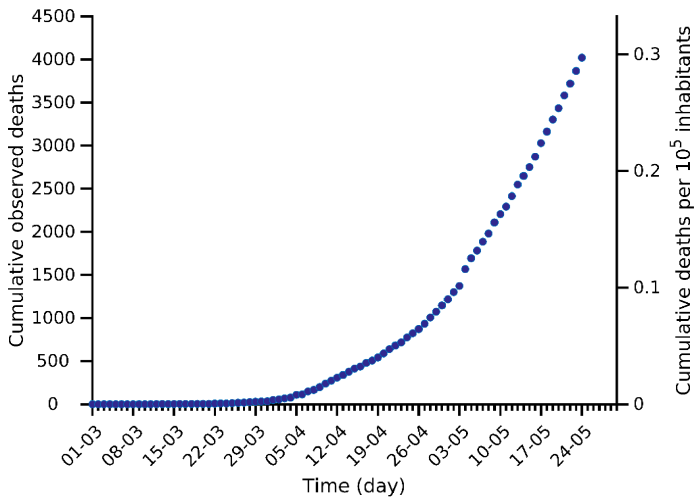


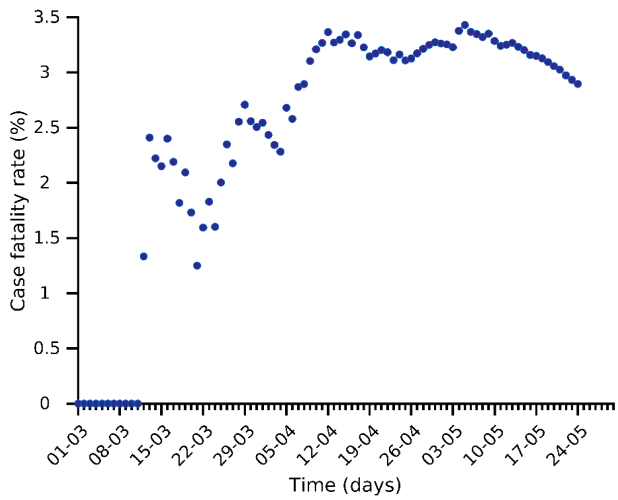




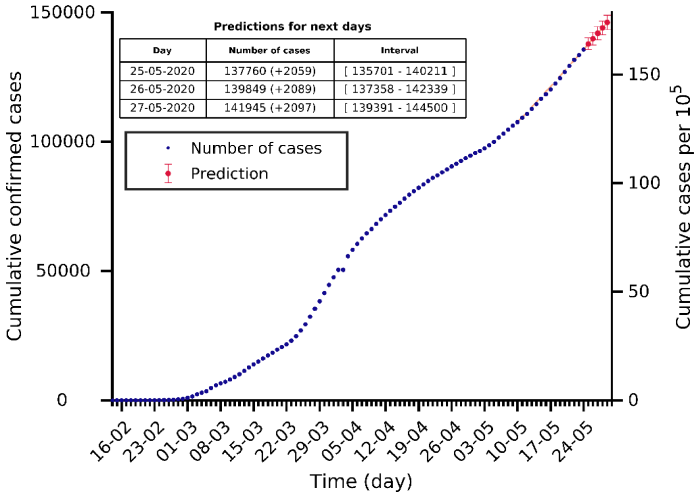


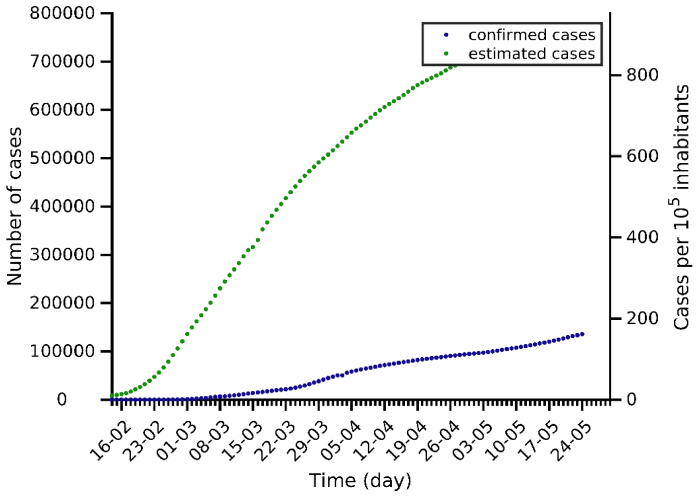


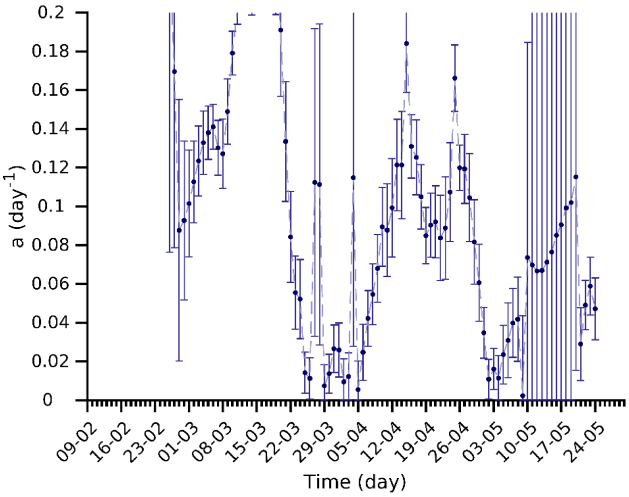


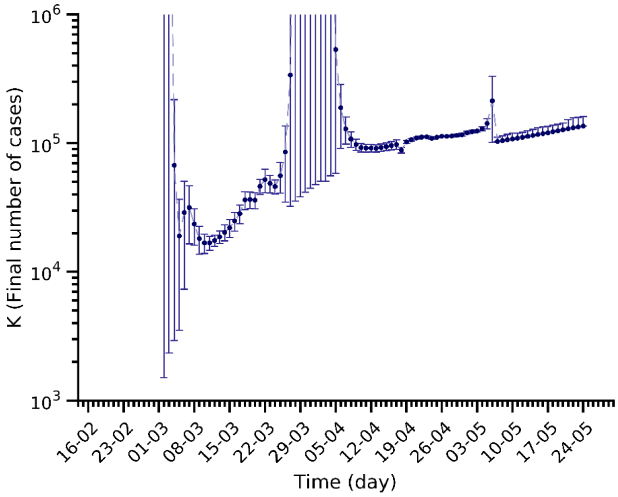


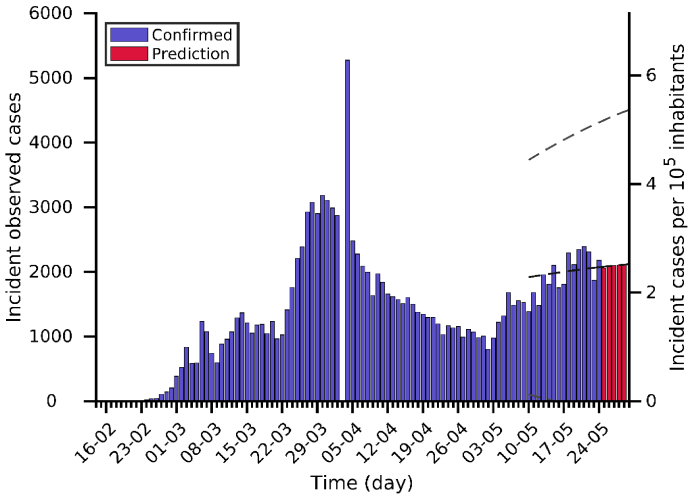
Iran 24-05-2020. Population: 84.0M. Current cumulated incidence: 162/10^5

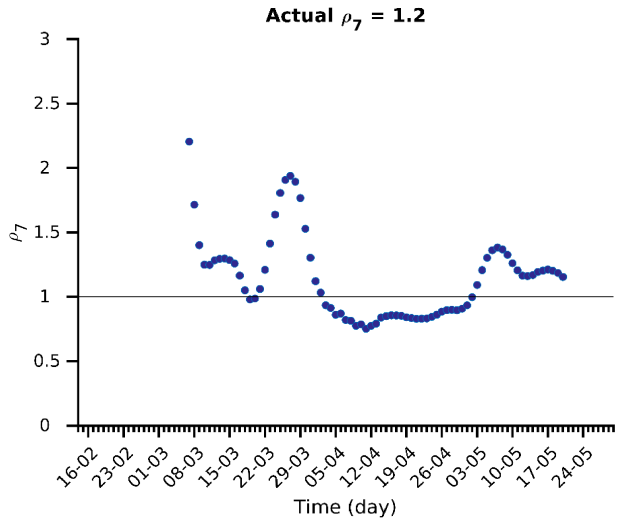


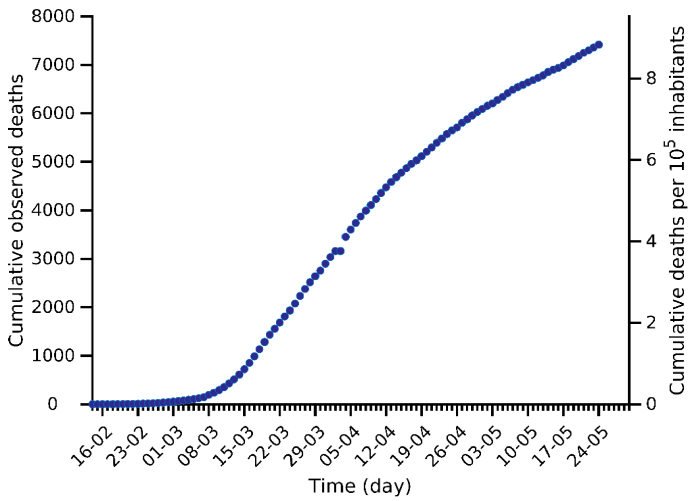


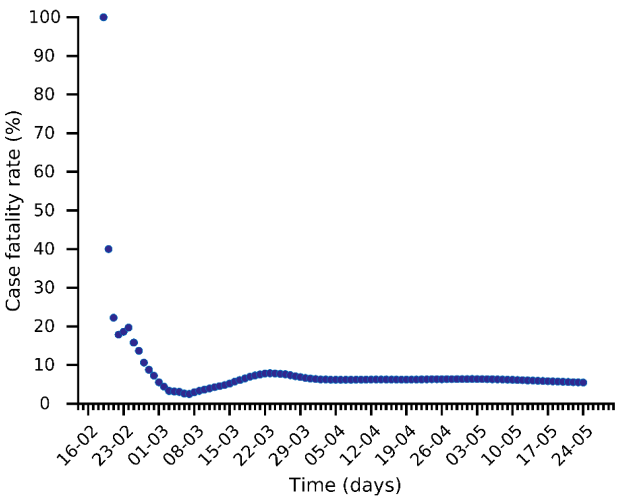




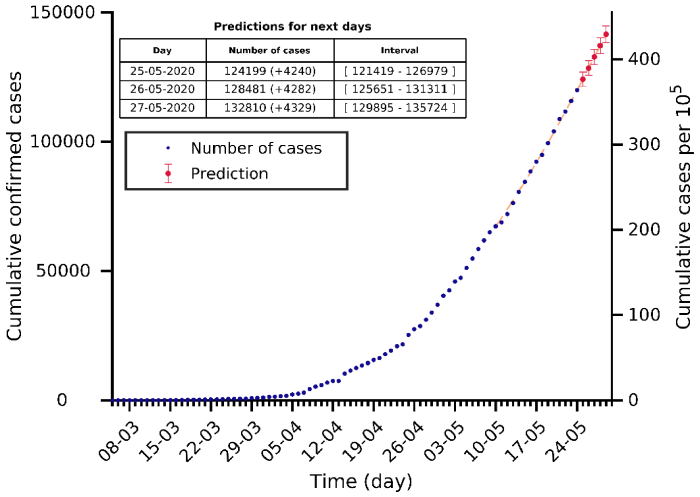


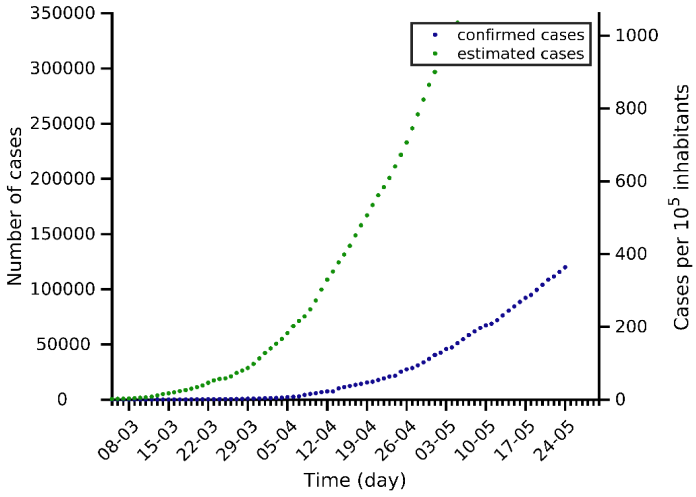


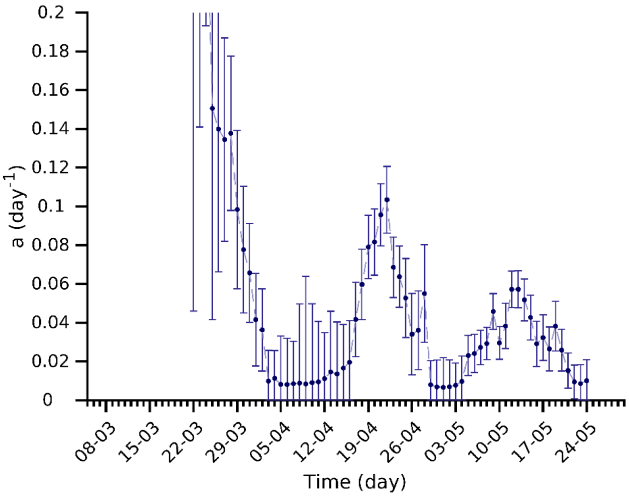


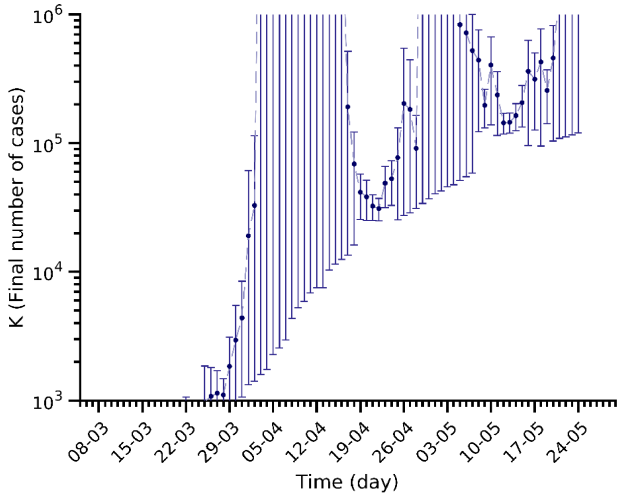


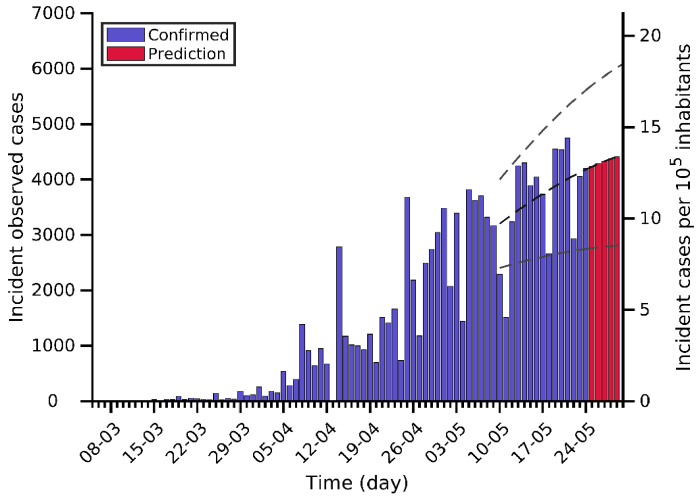
Peru 24-05-2020. Population: 33.0M. Current cumulated incidence: 364/10^5

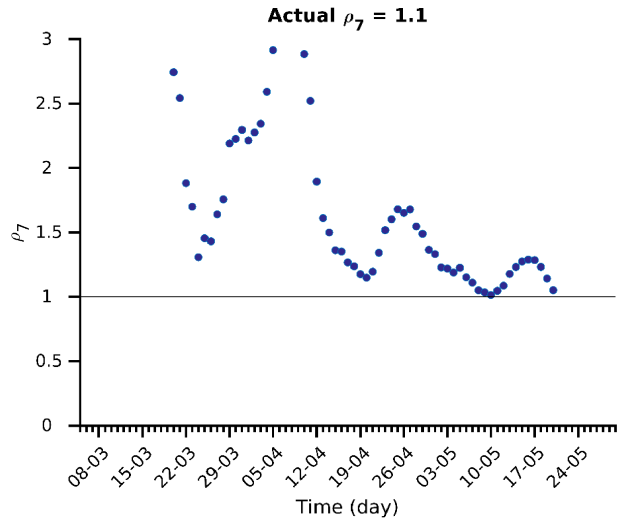


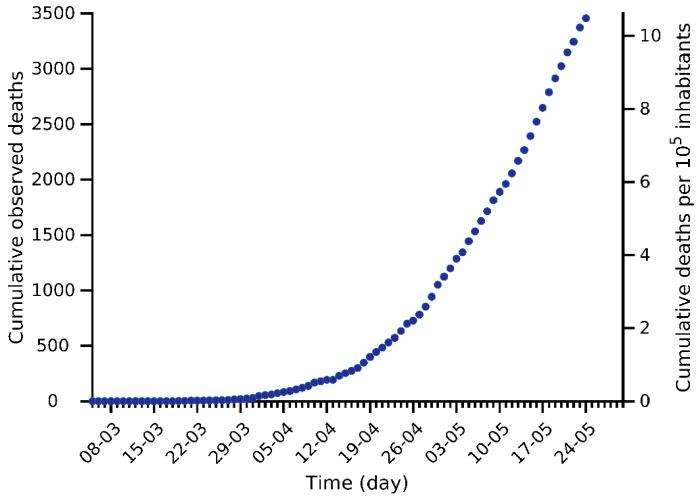


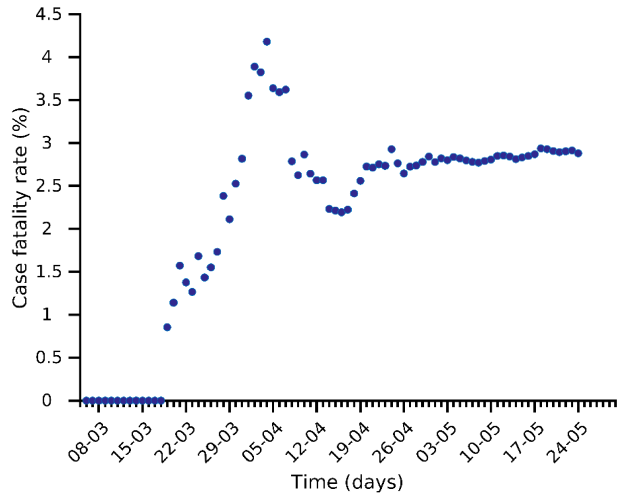




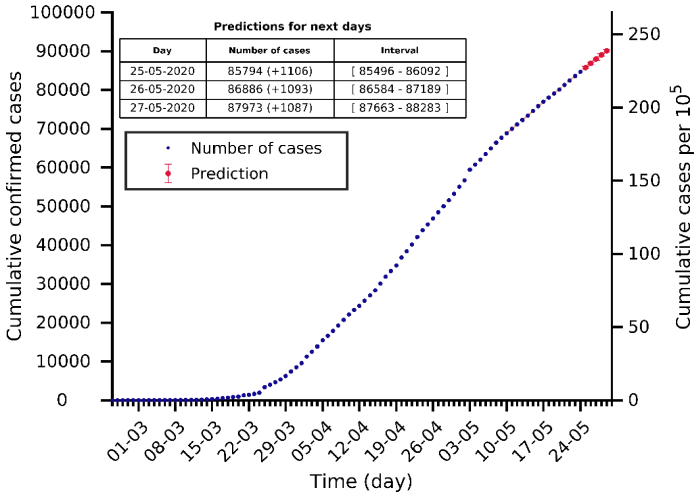


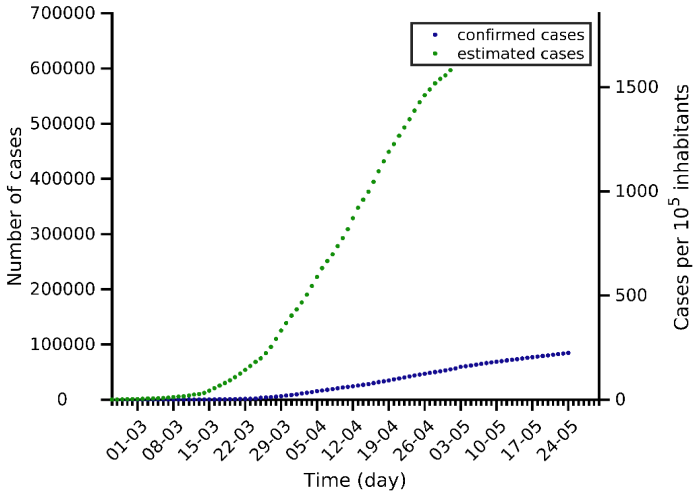


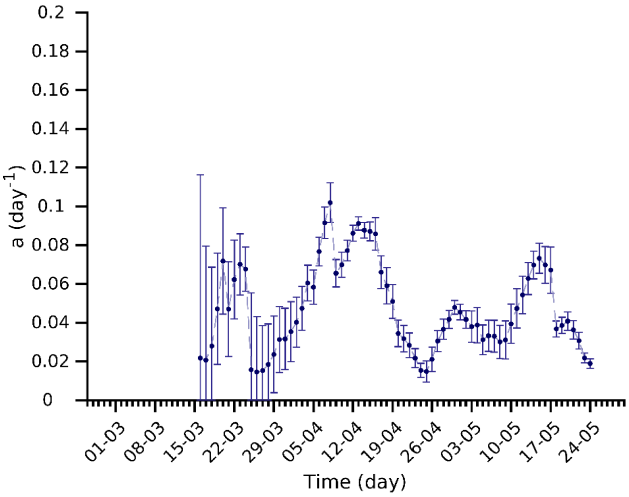


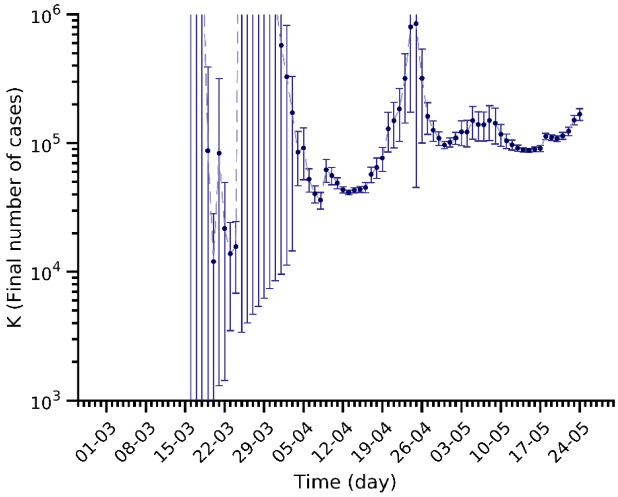


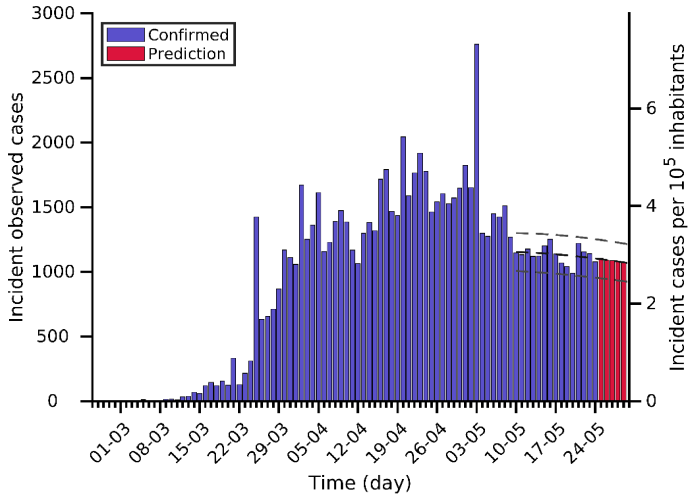
Canada 24-05-2020. Population: 37.7M. Current cumulated incidence: 224/10^5

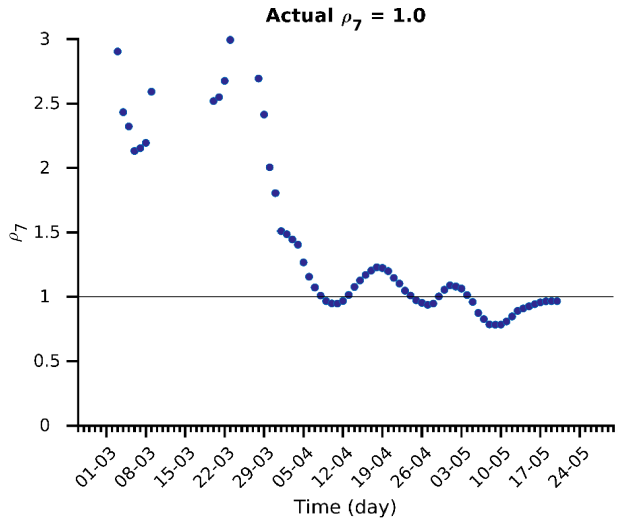


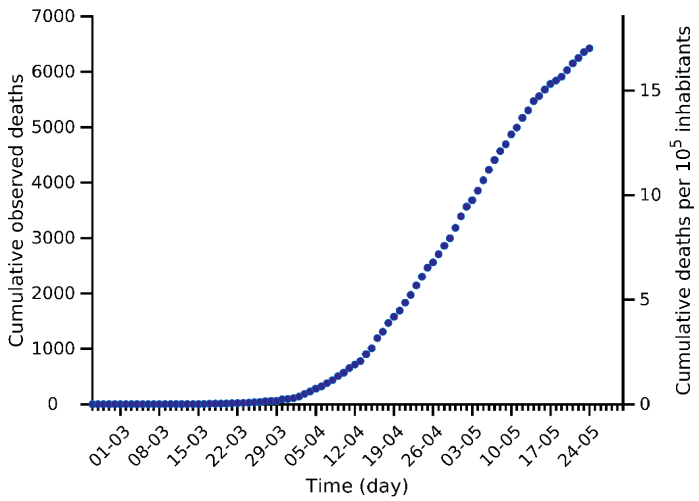


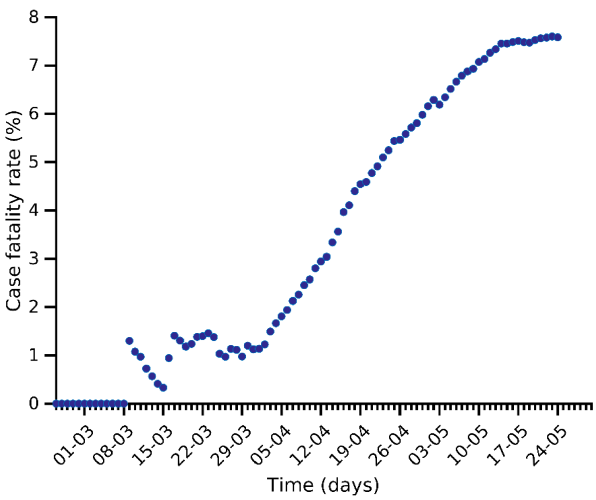




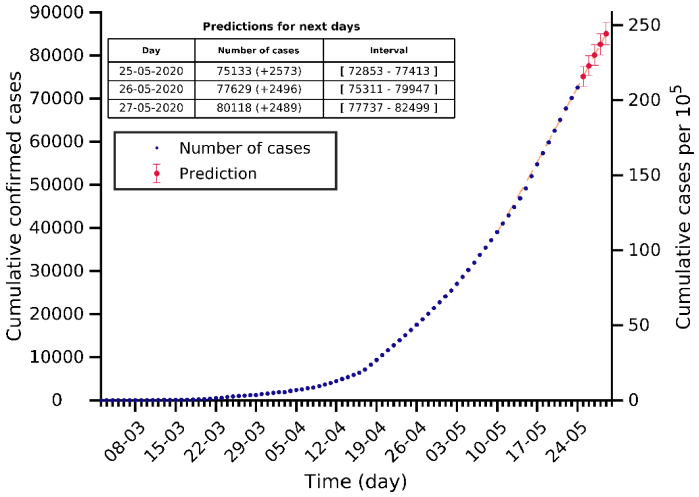


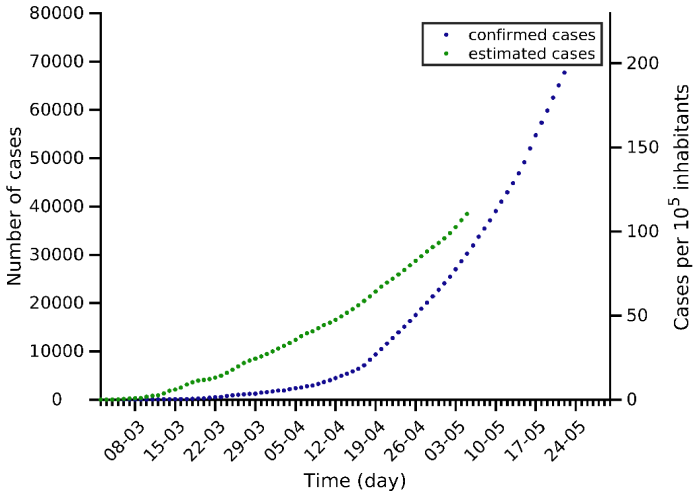


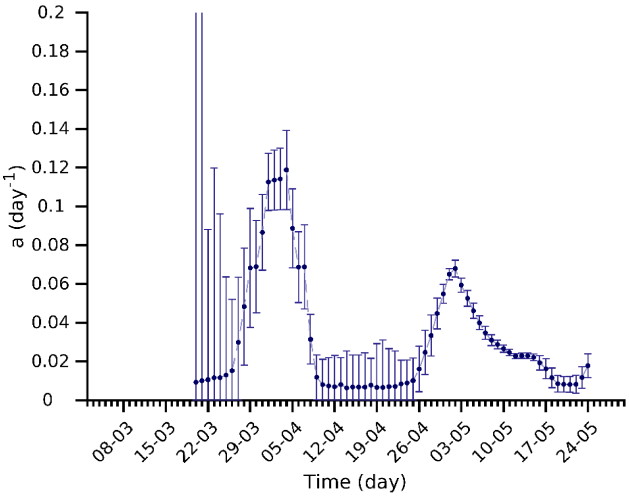


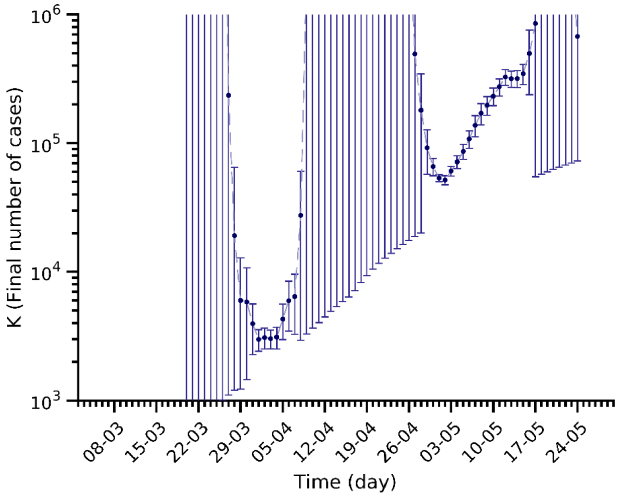


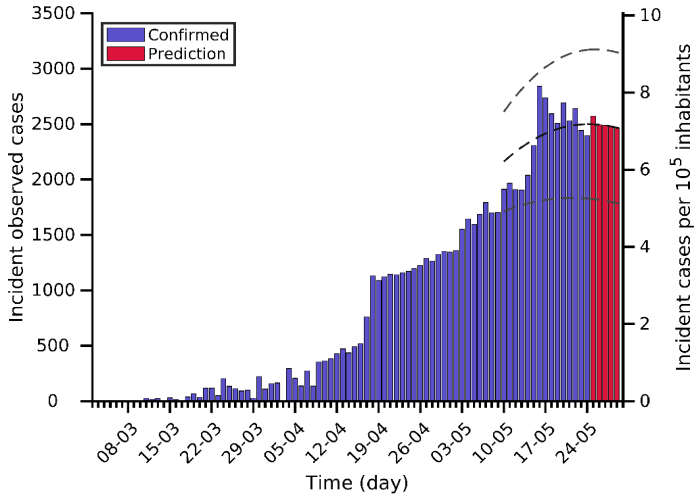
Saudi Arabia 24-05-2020. Population: 34.8M. Current cumulated incidence: 208/10^5

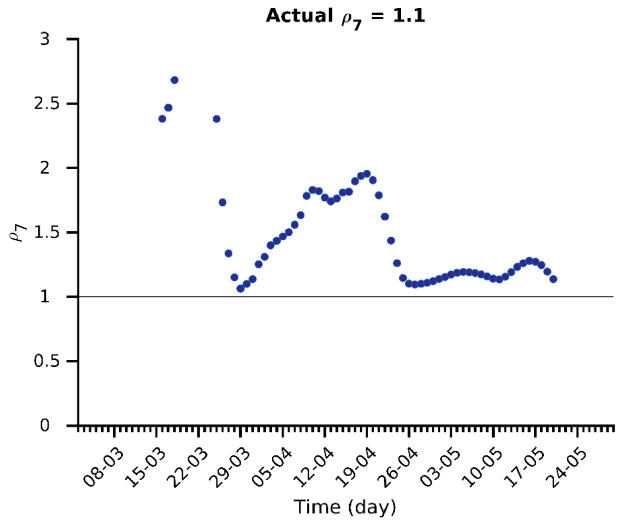


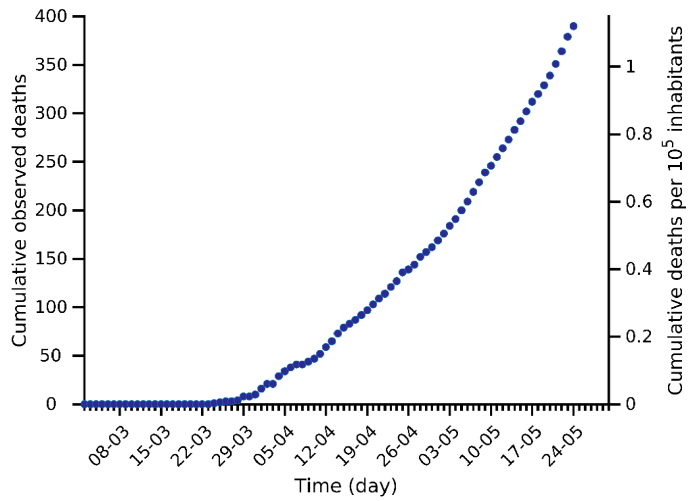


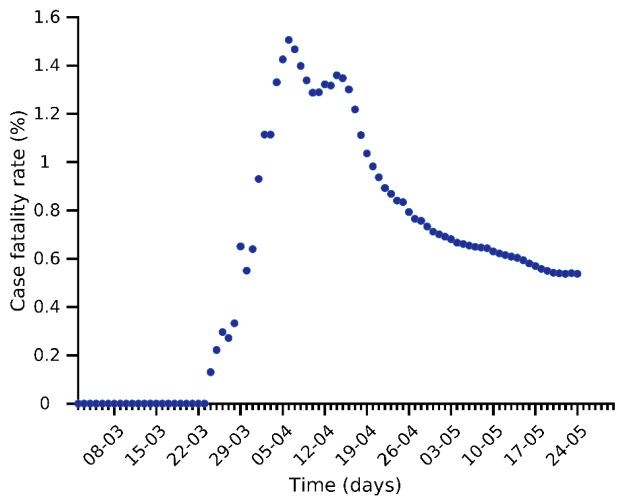




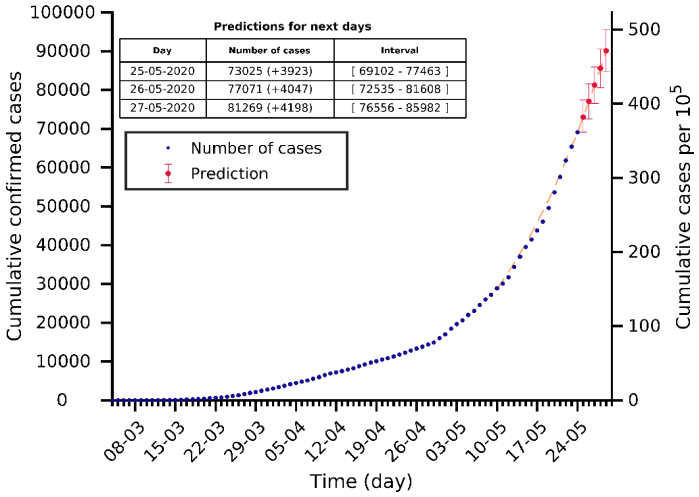


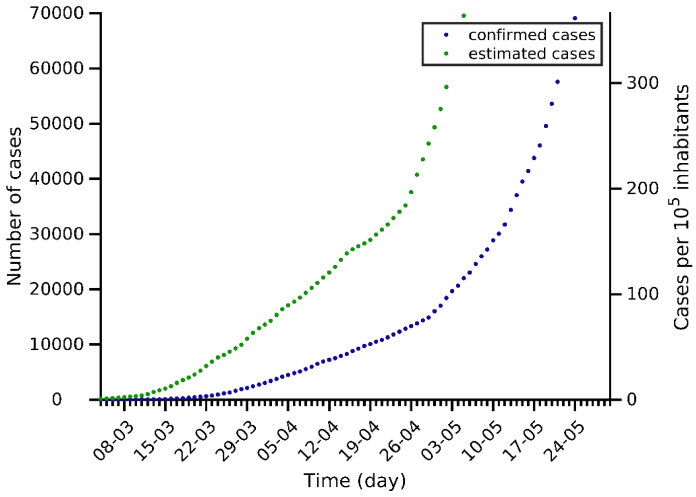


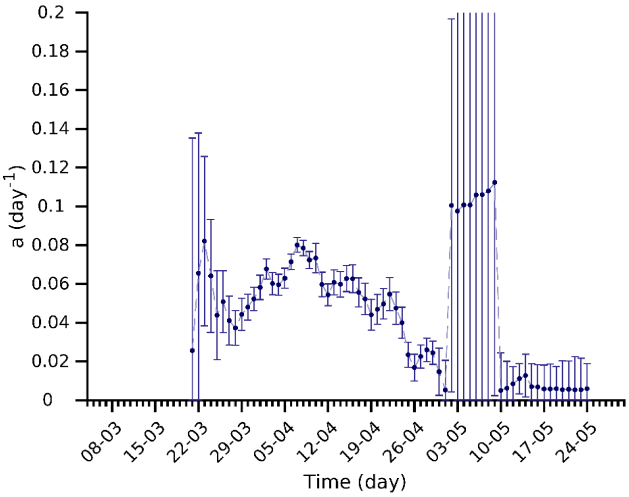


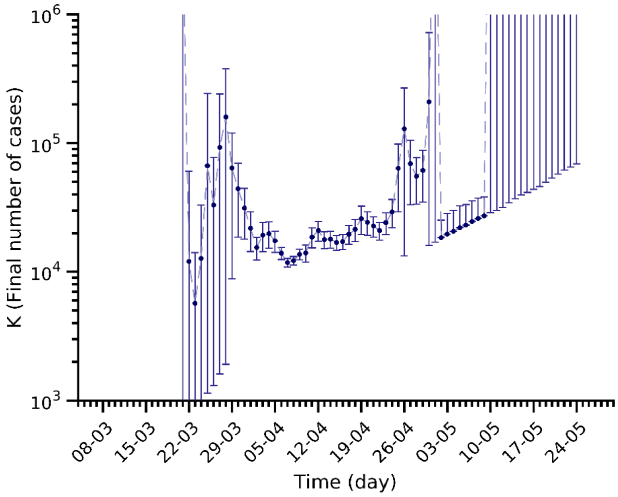


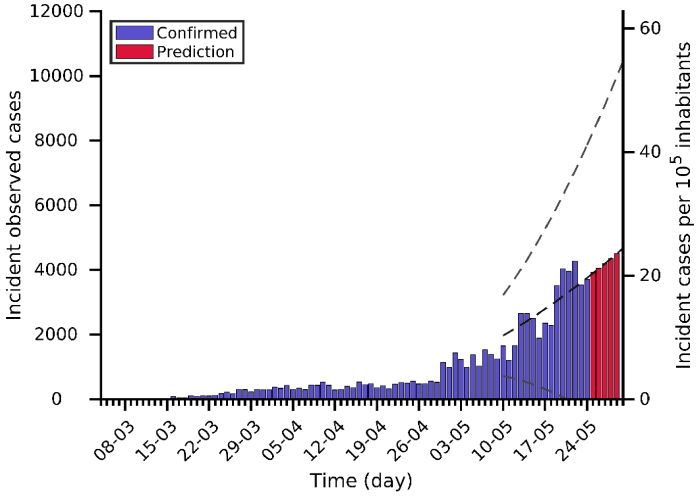
Chile 24-05-2020. Population: 19.1M. Current cumulated incidence: 361/10^5

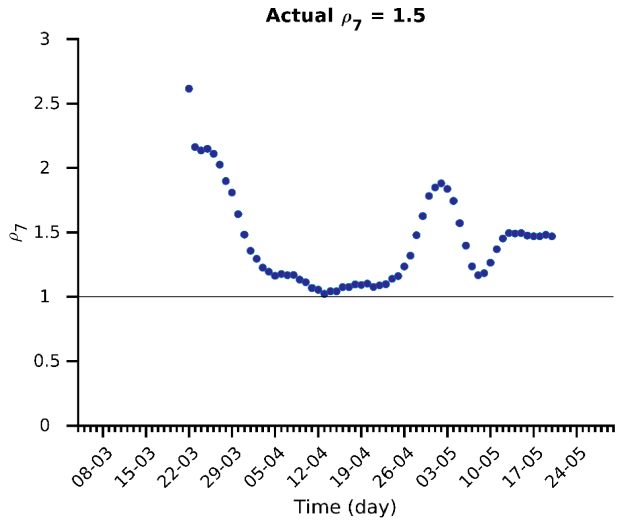


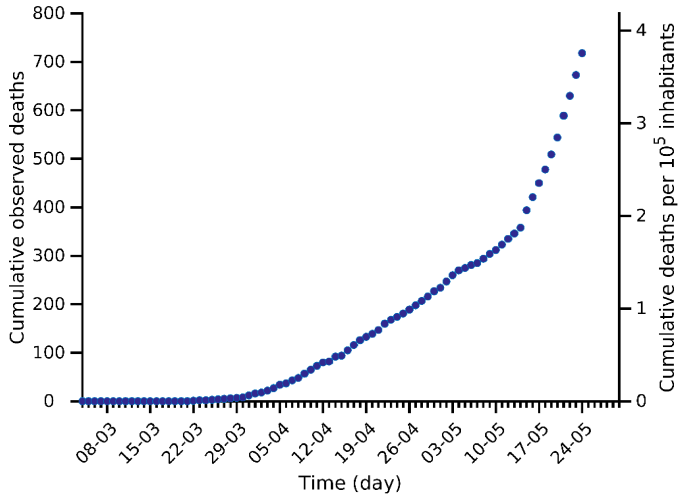


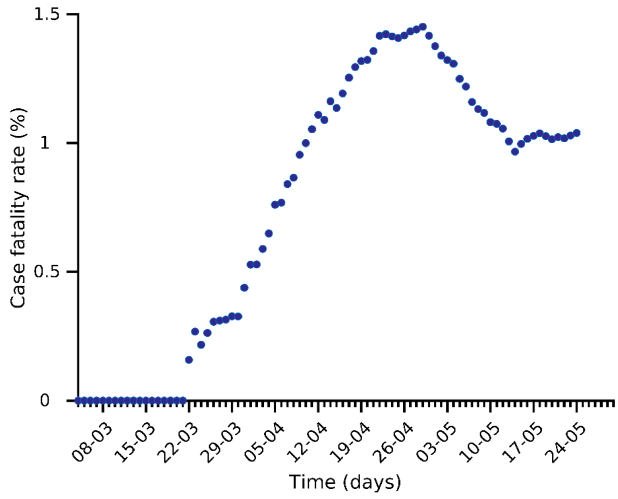




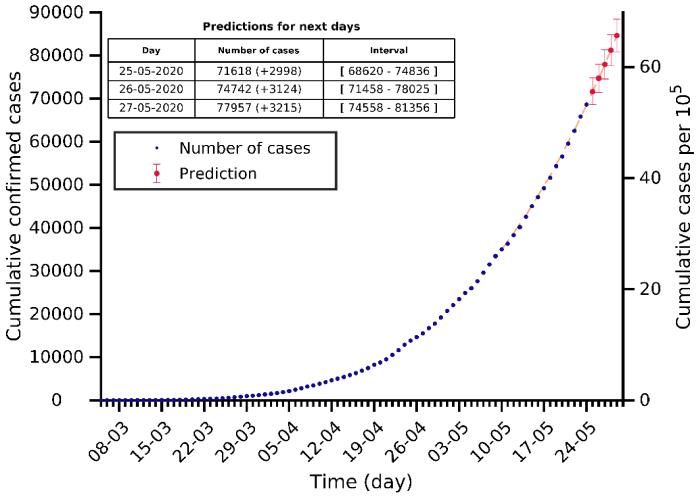


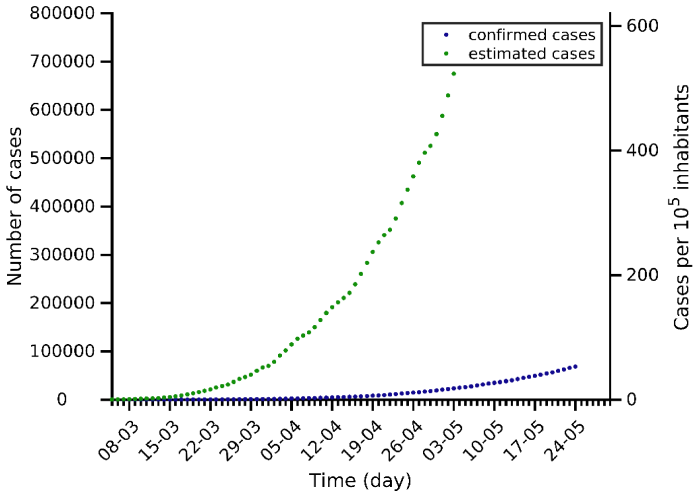


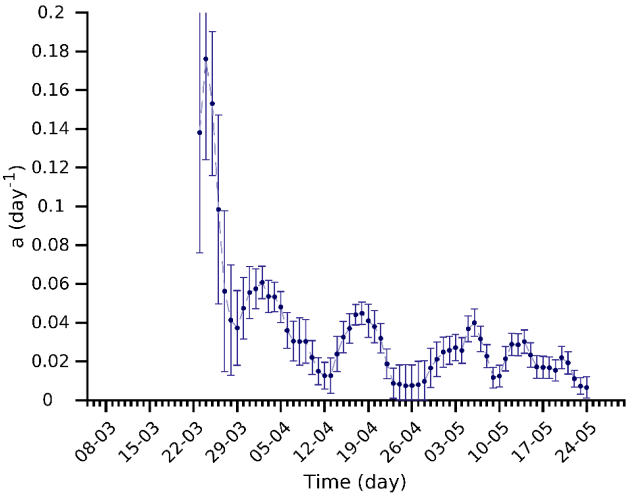


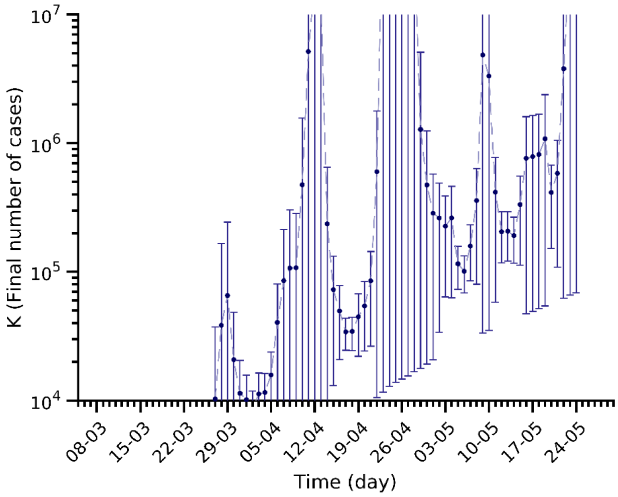


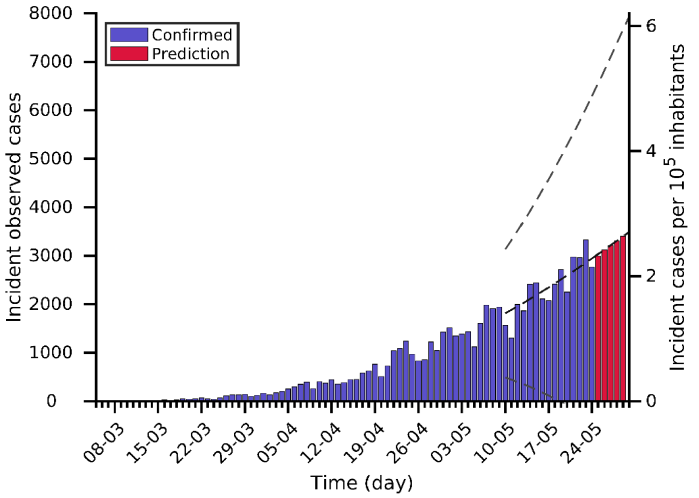
Mexico 24-05-2020. Population: 128.9M. Current cumulated incidence: 53/10^5

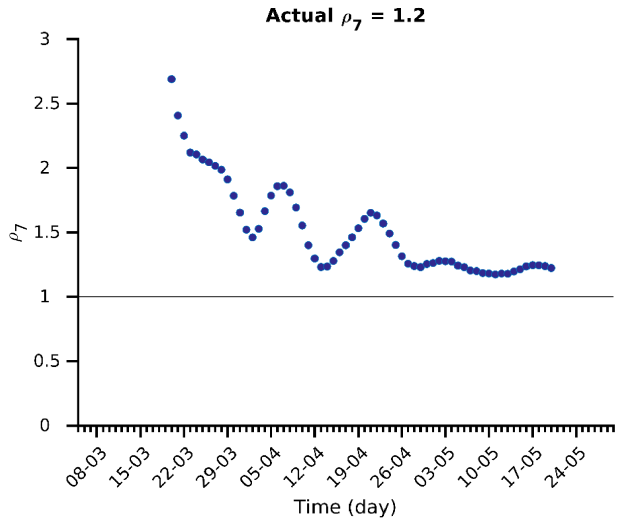


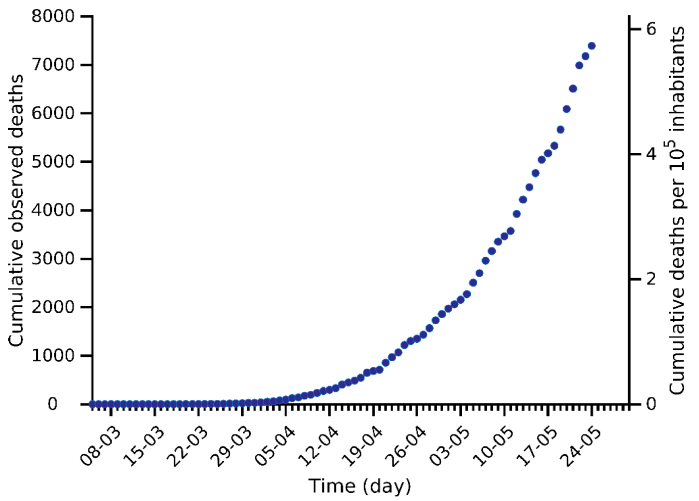


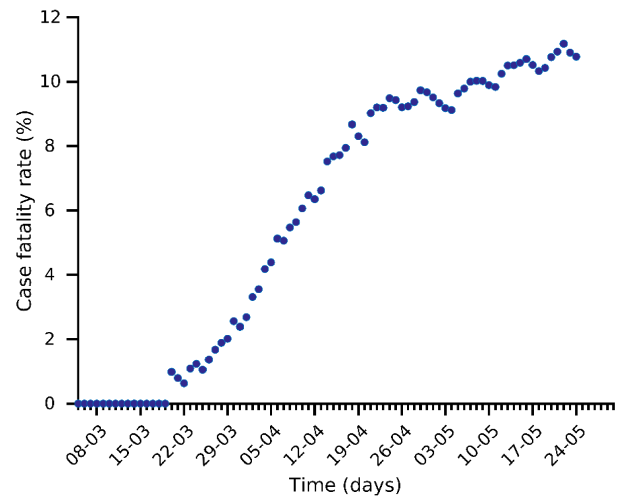




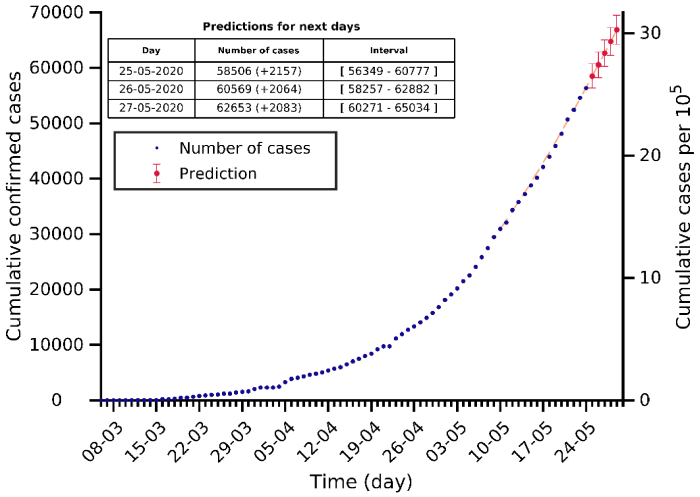


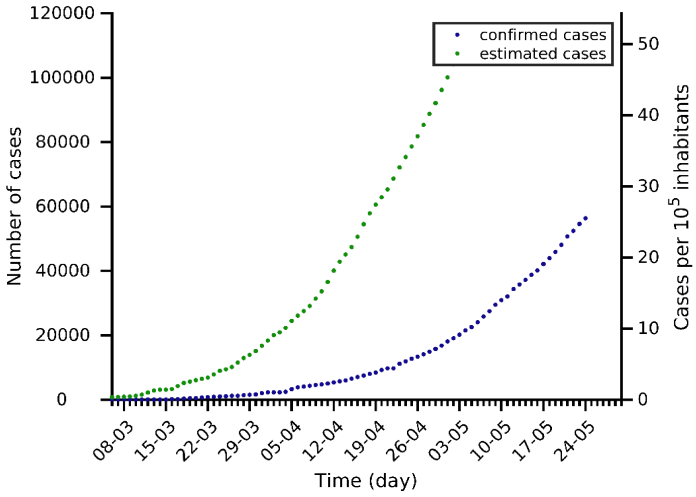


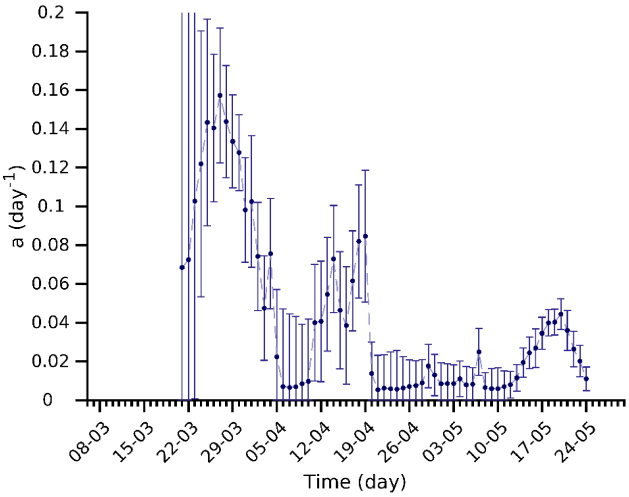


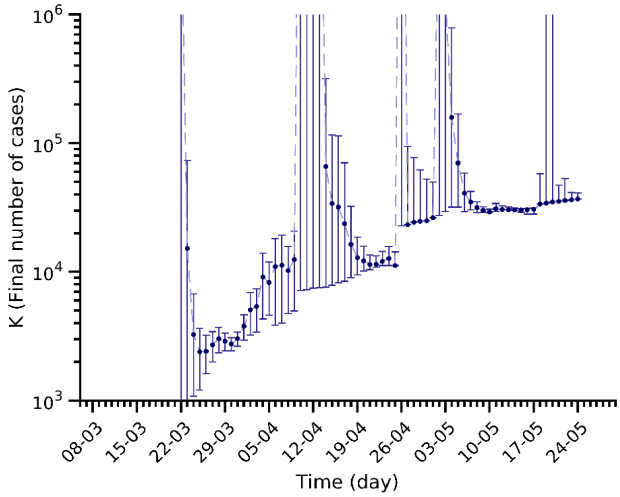


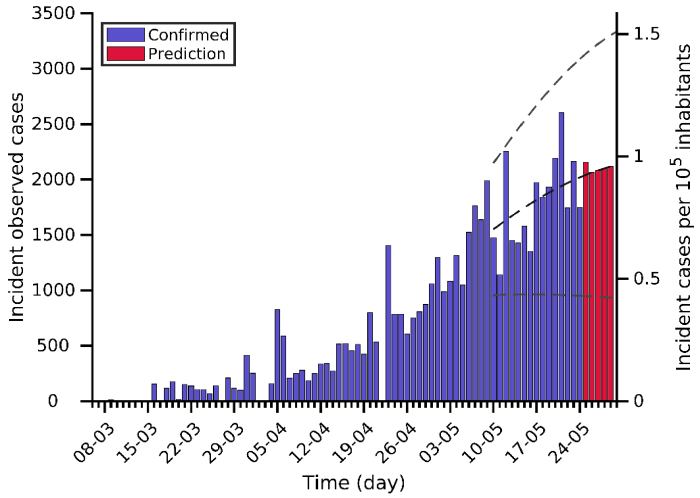
Pakistan 24-05-2020. Population: 220.9M. Current cumulated incidence: $26/10^5$

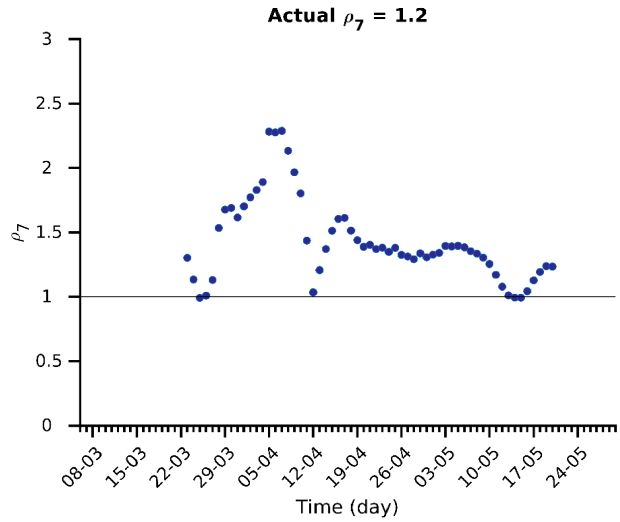


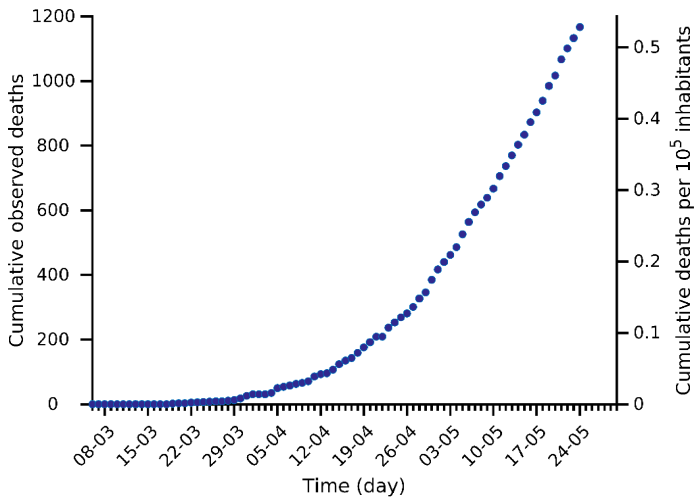


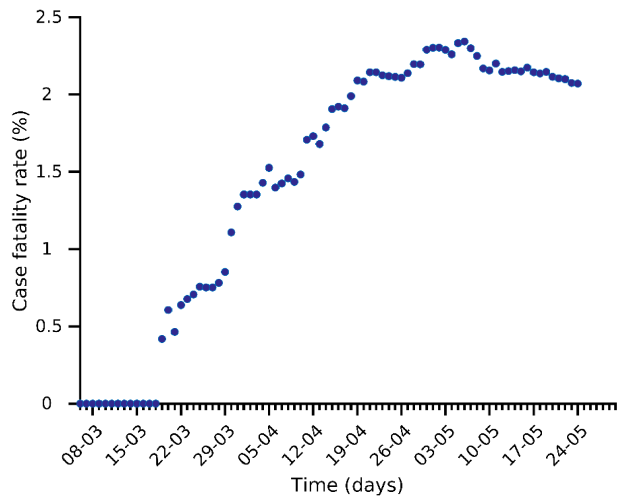




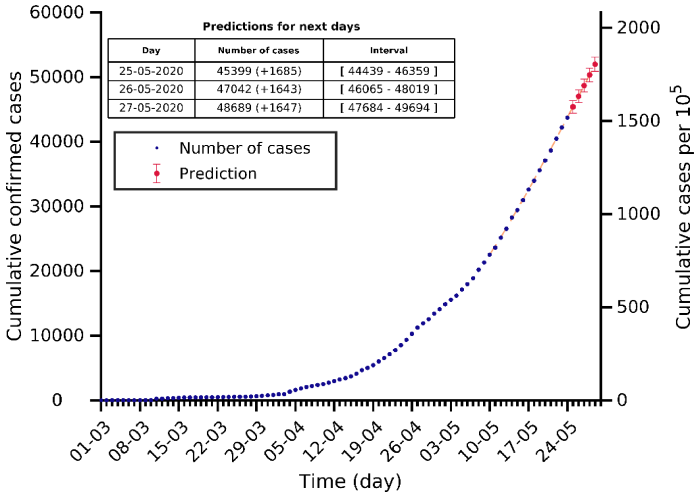


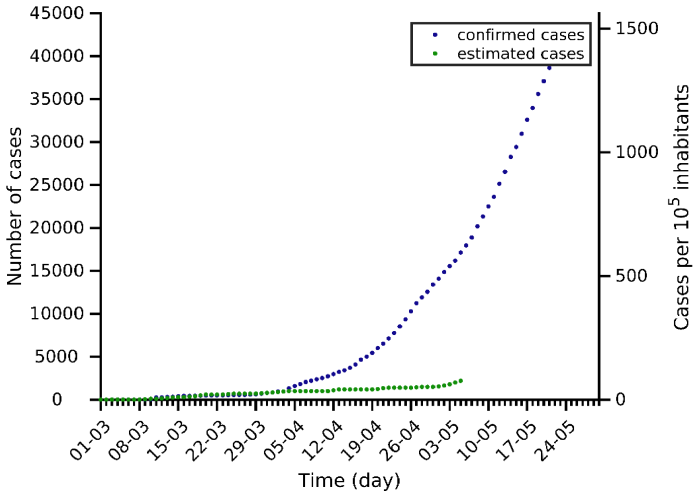


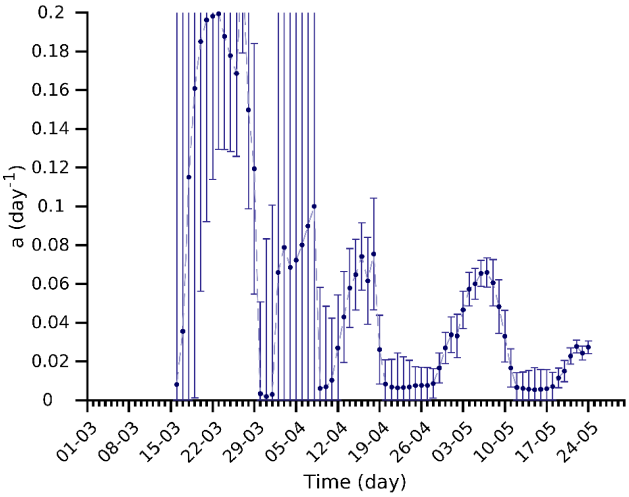


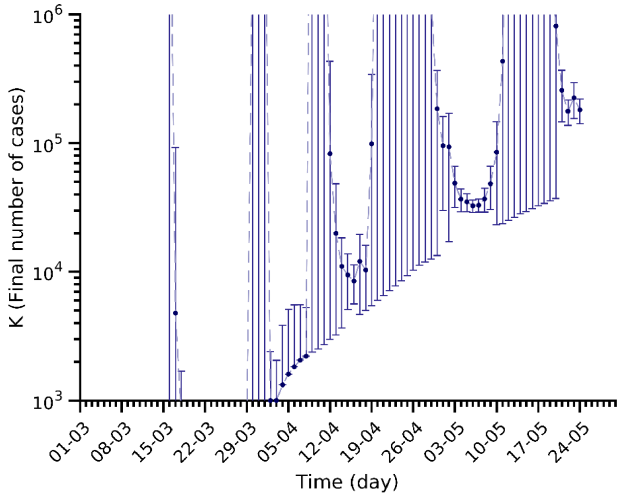


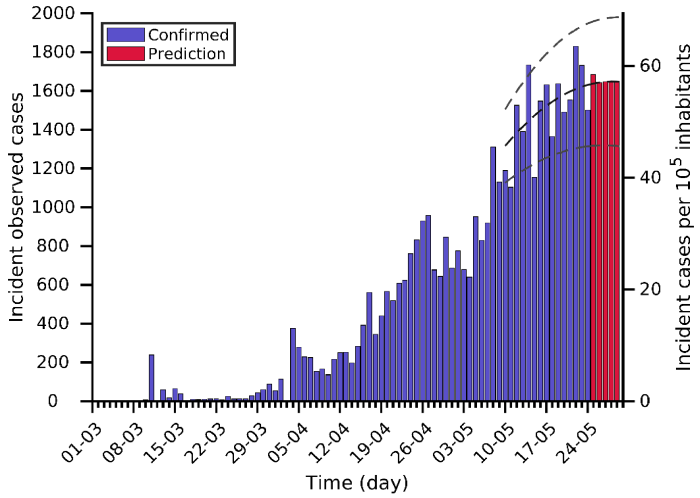
Qatar 24-05-2020. Population: 2.9M. Current cumulated incidence: 1517/10⁵

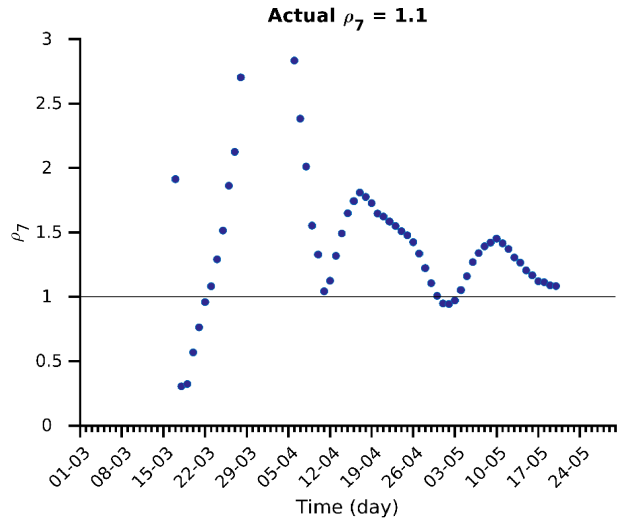


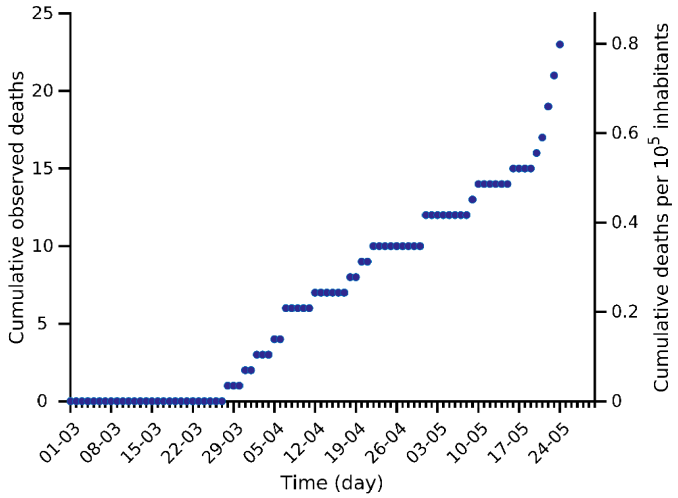


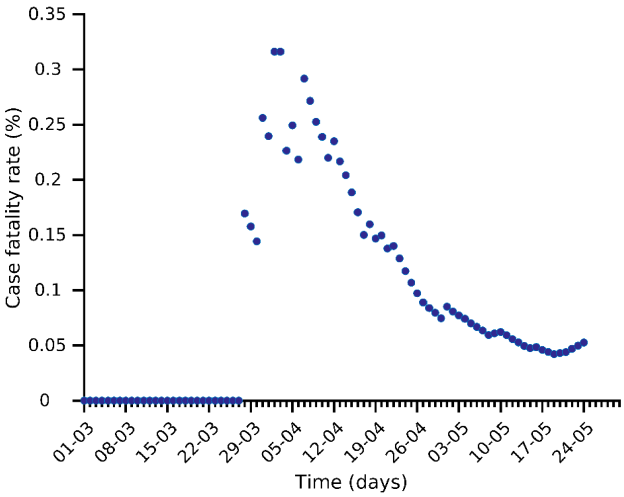




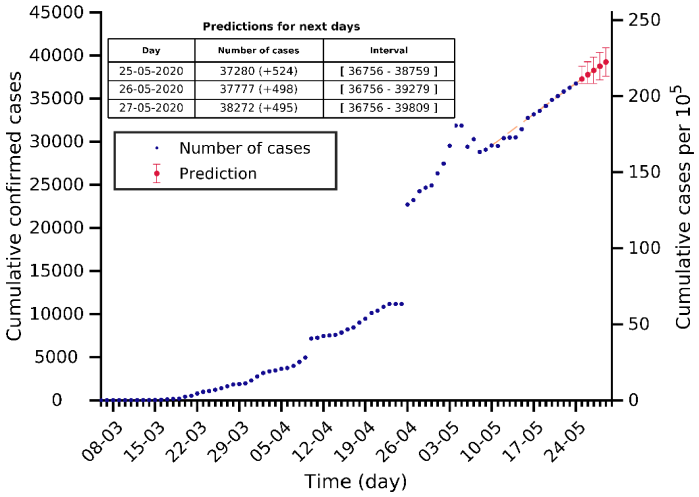


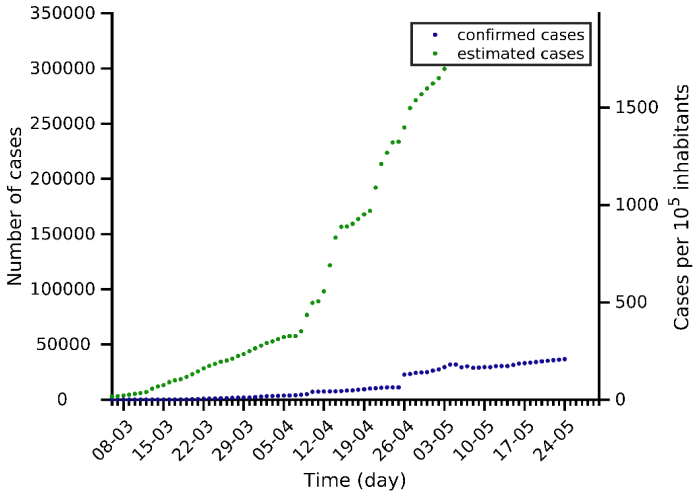


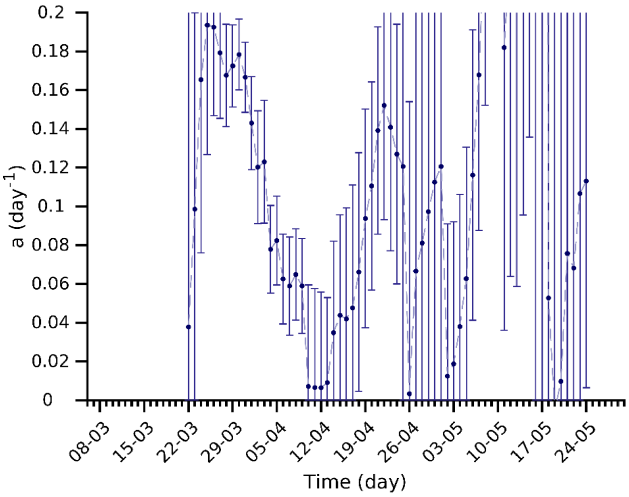


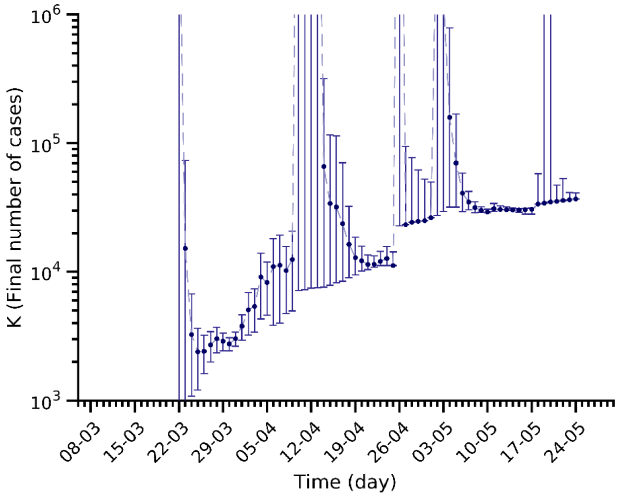


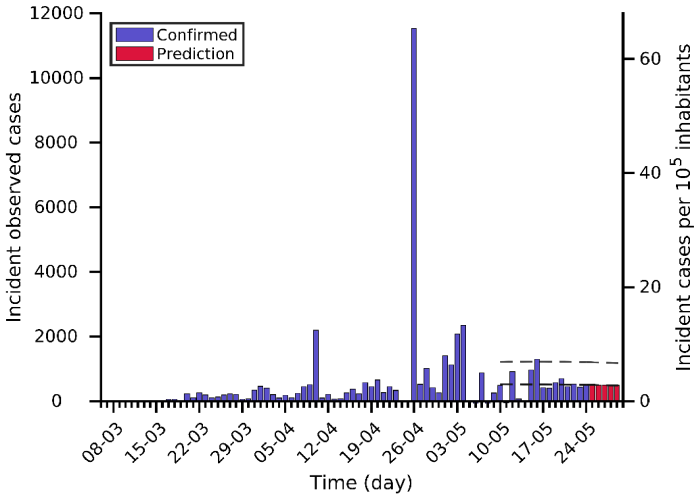
Ecuador 24-05-2020. Population: 17.6M. Current cumulated incidence: 208/10^5

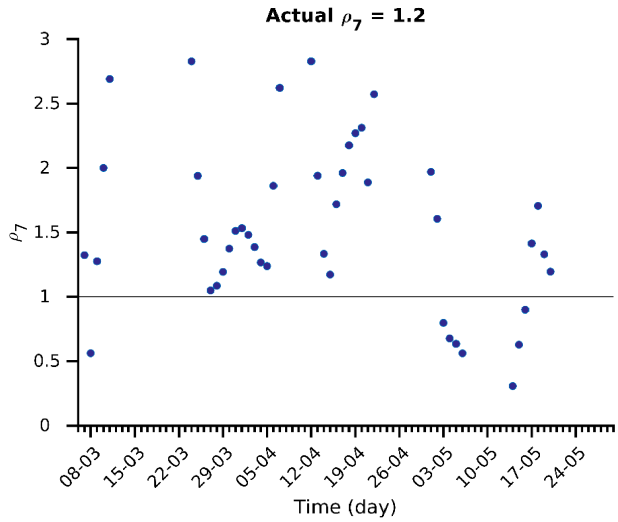


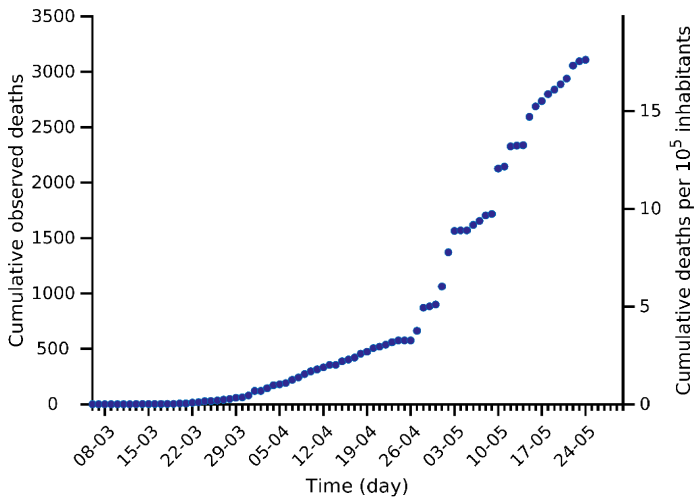


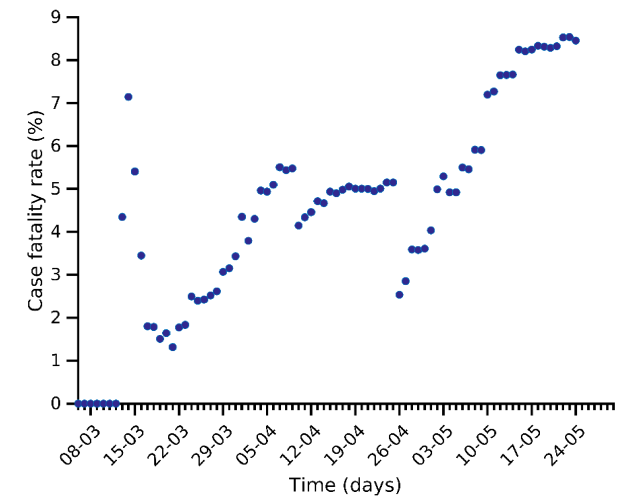




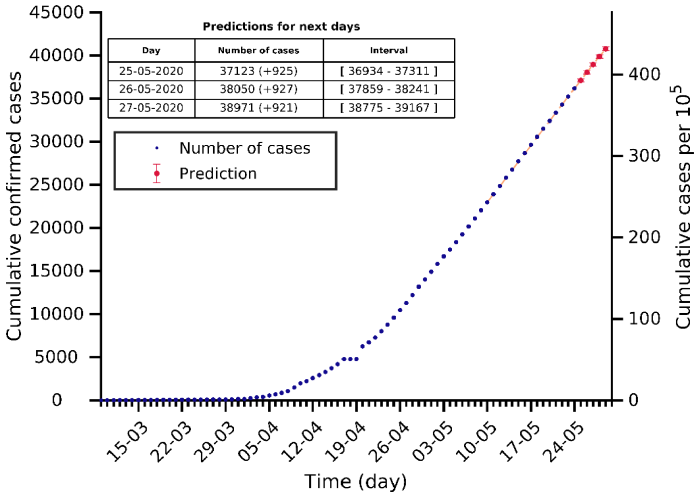


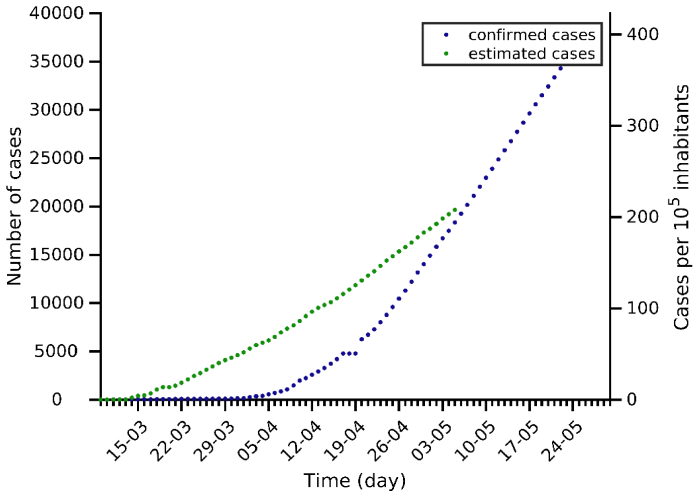


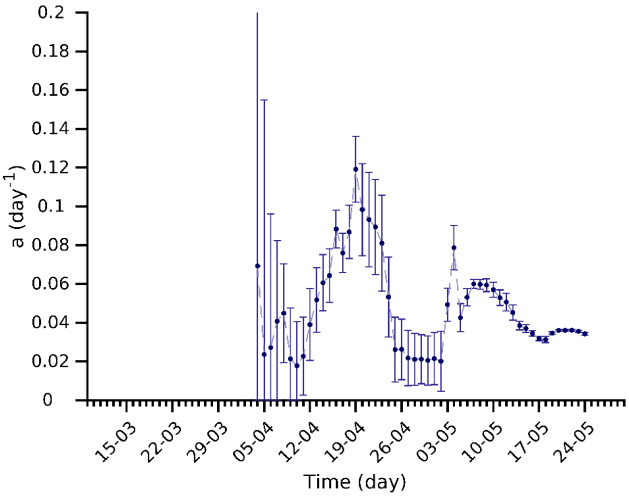


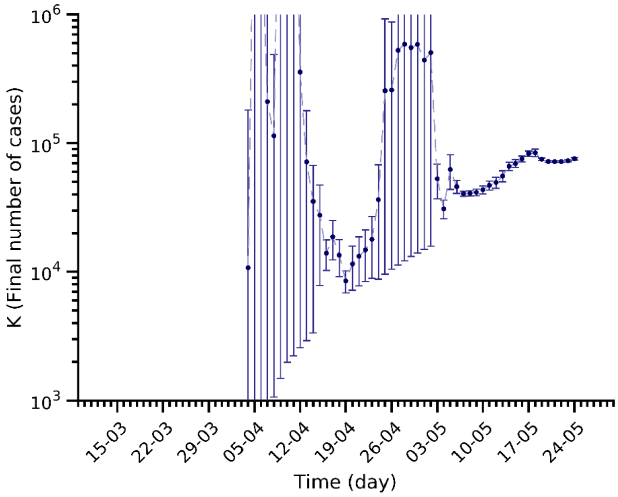


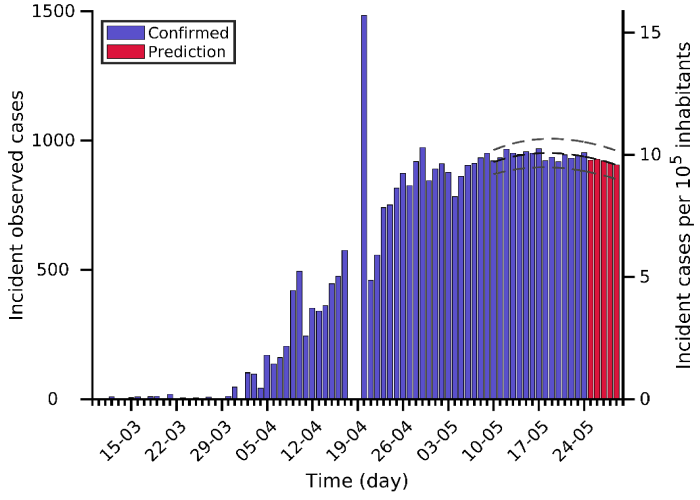
Belarus 24-05-2020. Population: 9.4M. Current cumulated incidence: 383/10^5

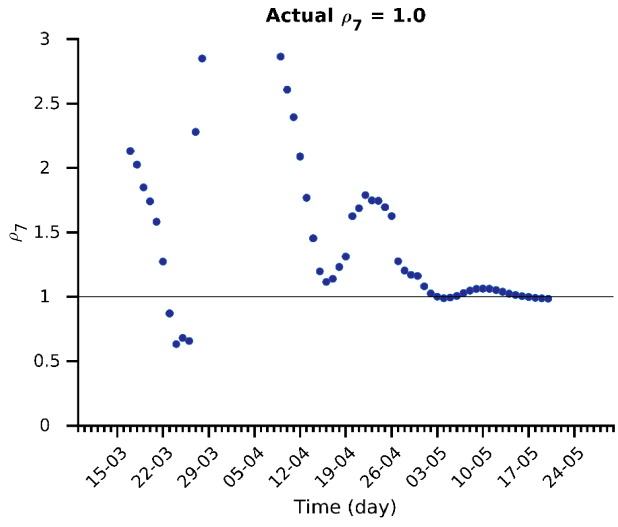


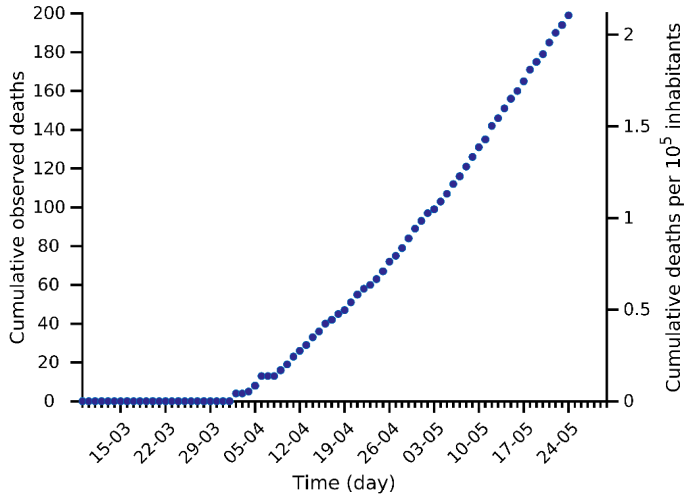


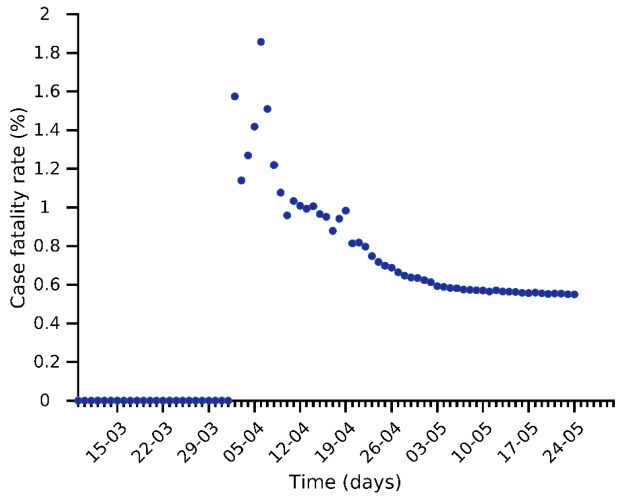




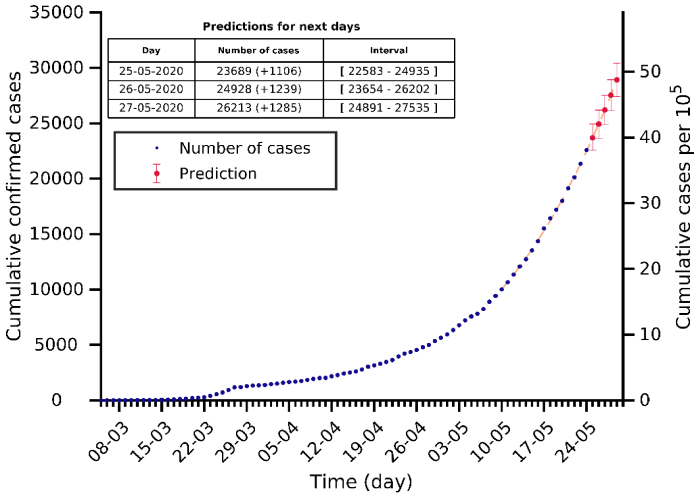


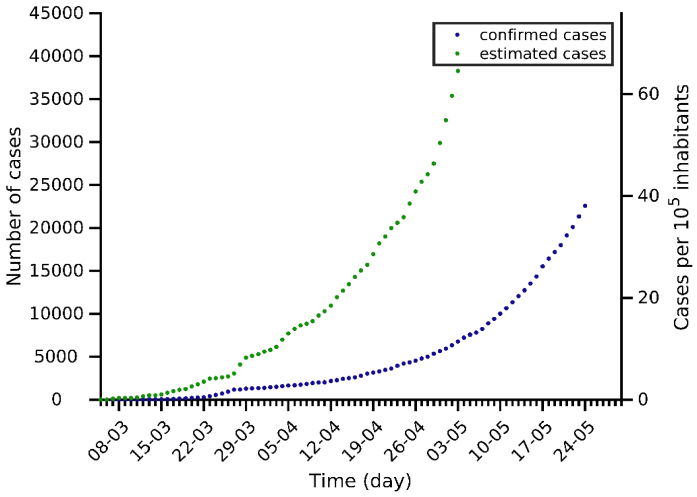


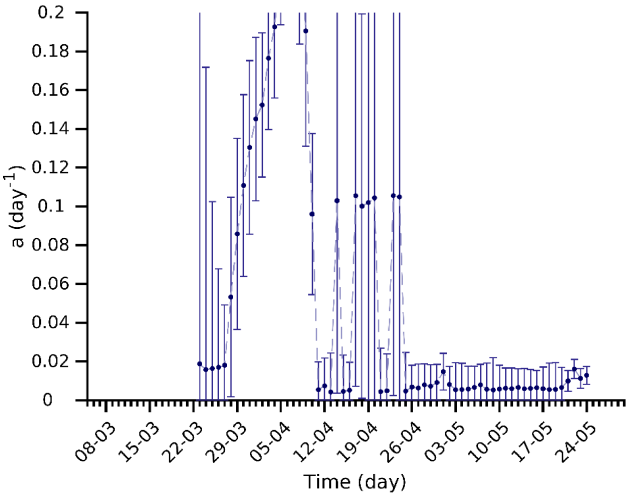


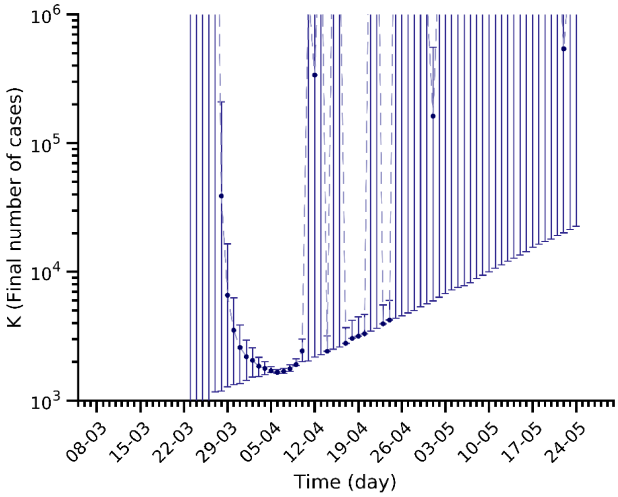


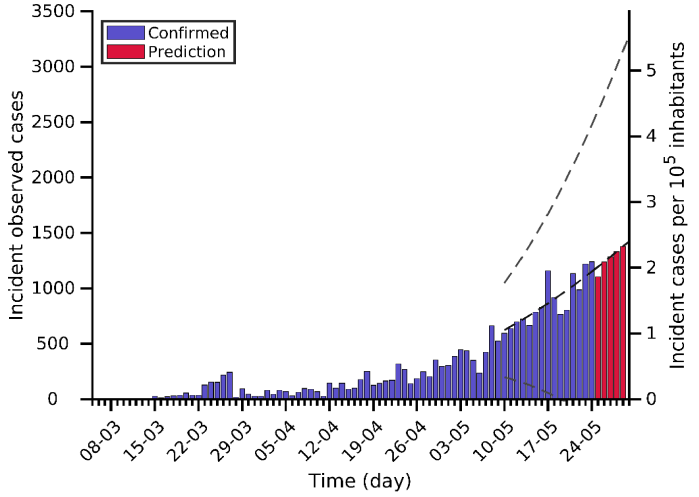
South Africa 24-05-2020. Population: 59.3M. Current cumulated incidence: $38/10^5$

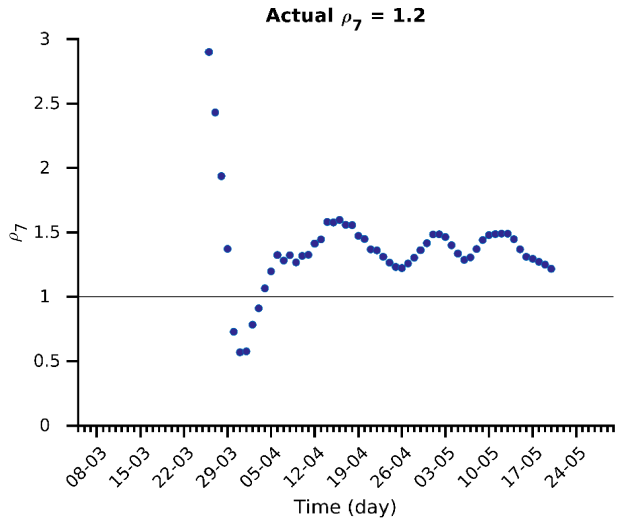


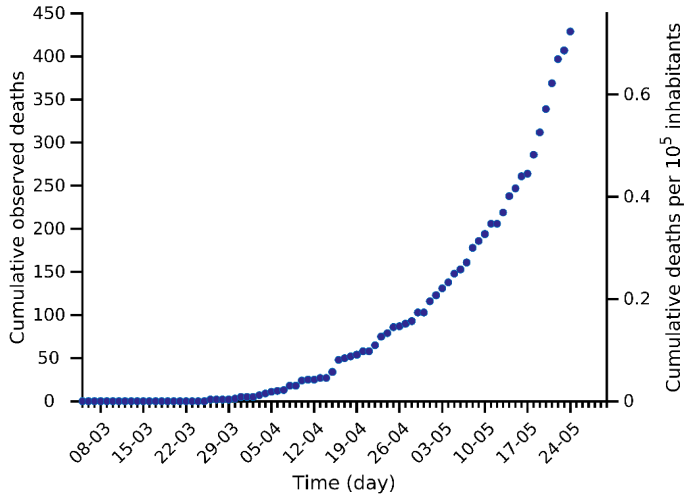


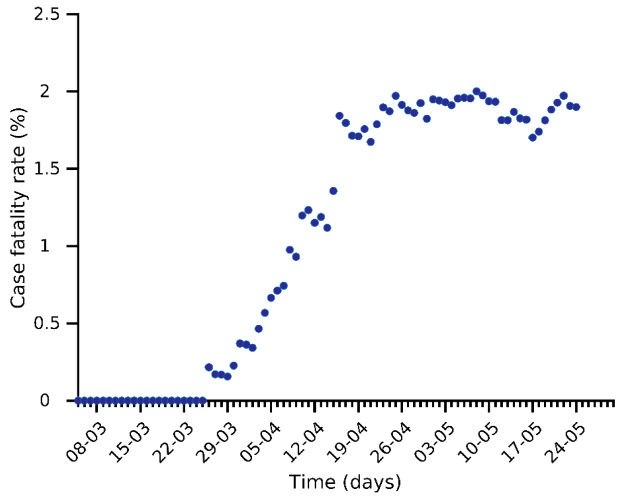




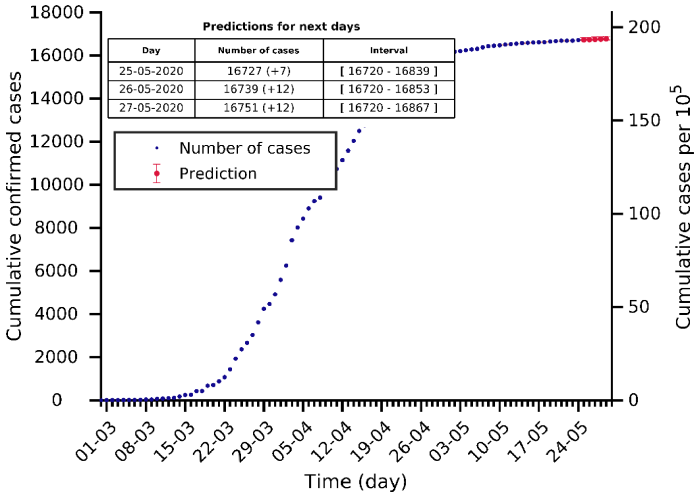


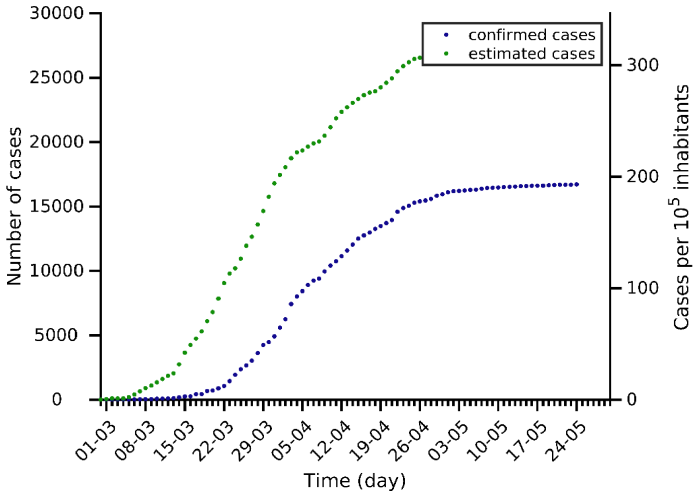


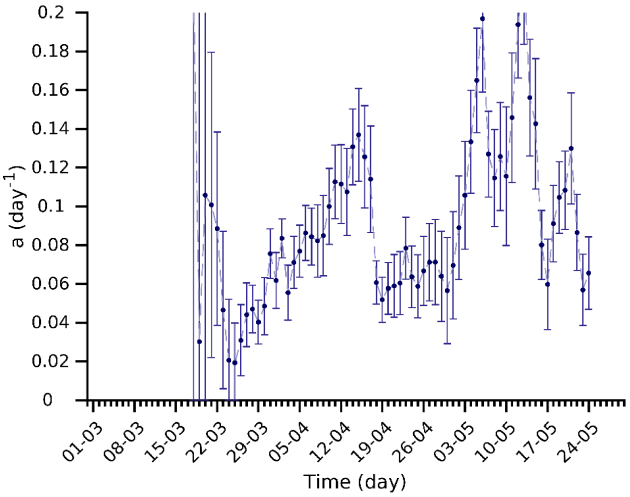


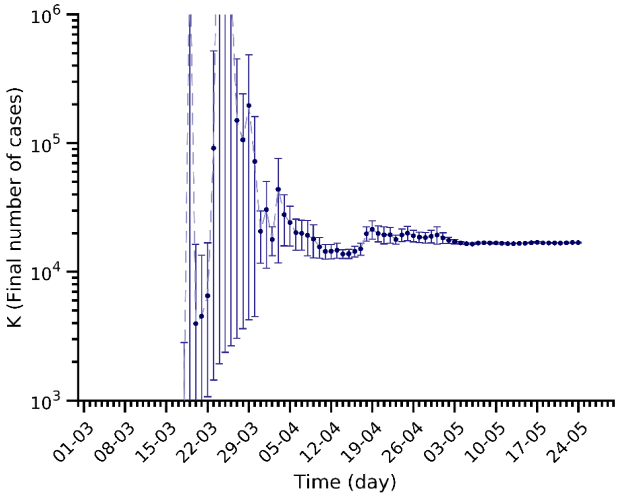


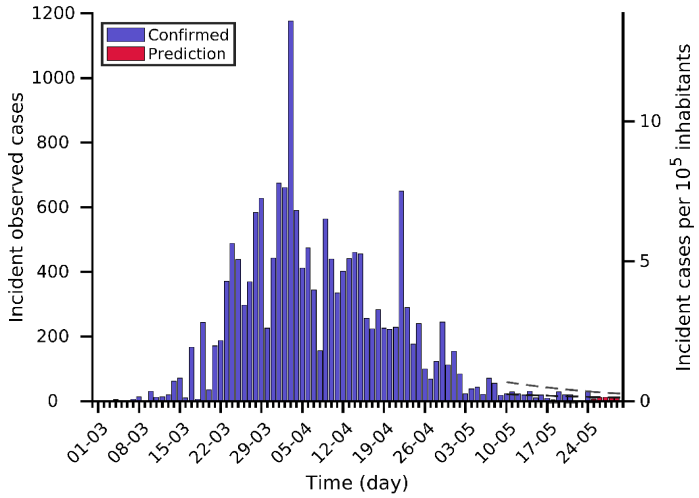
Israel 24-05-2020. Population: 8.7M. Current cumulated incidence: 193/10^5

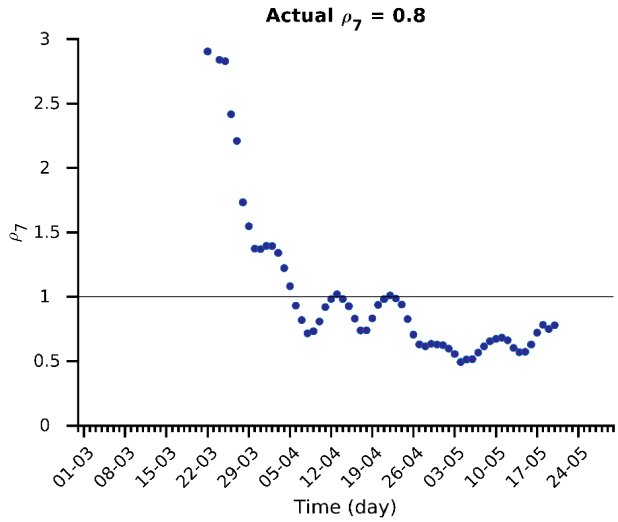


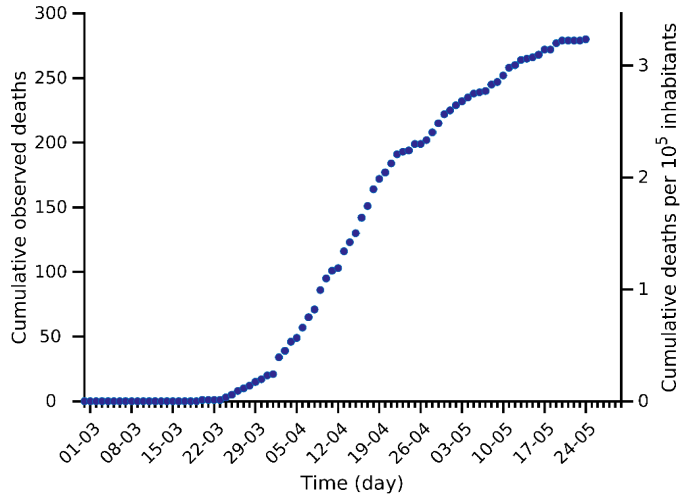


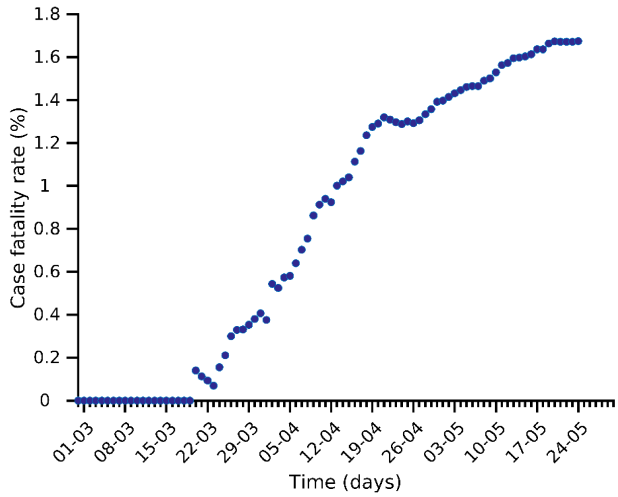




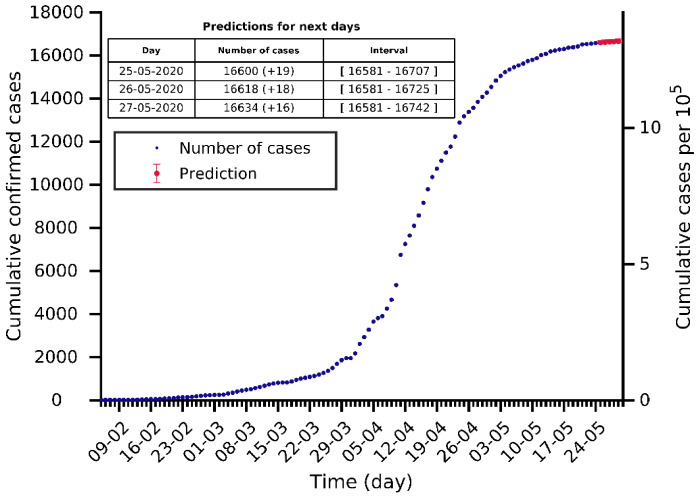


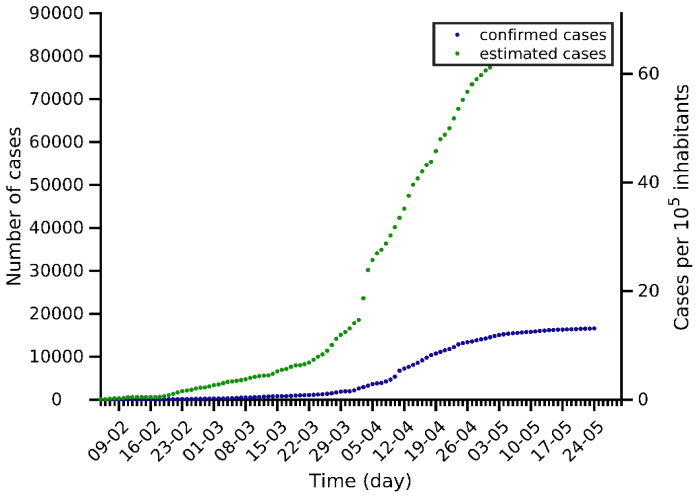


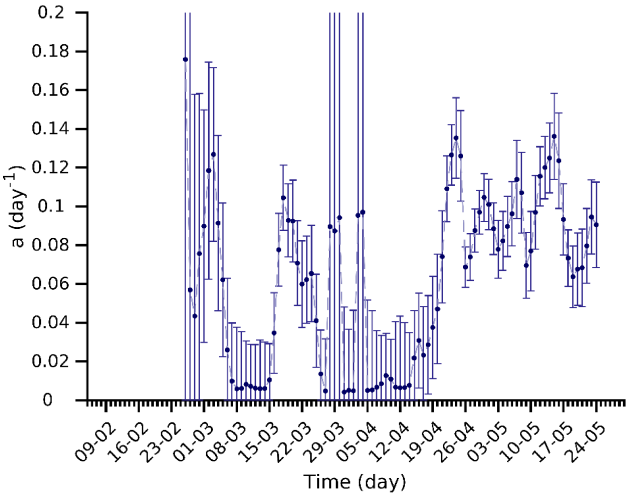


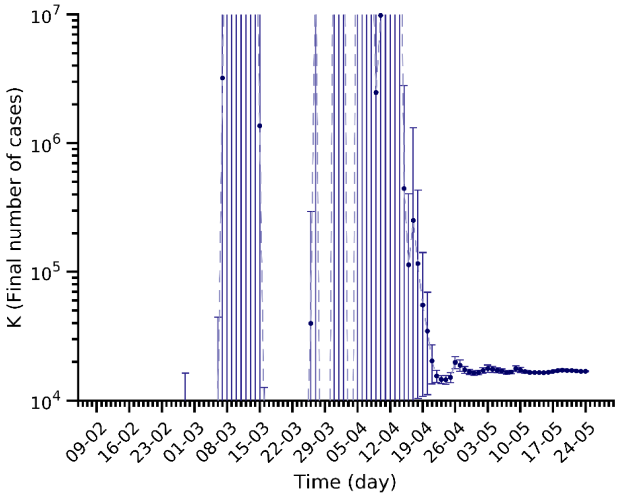


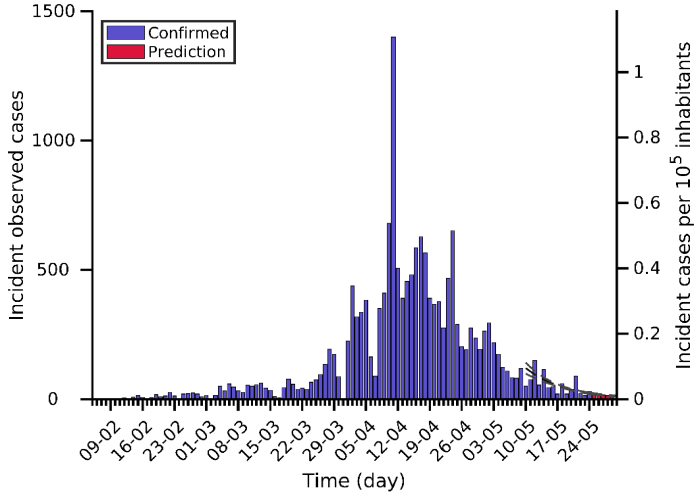
Japan 24-05-2020. Population: 126.5M. Current cumulated incidence: 13/10^5

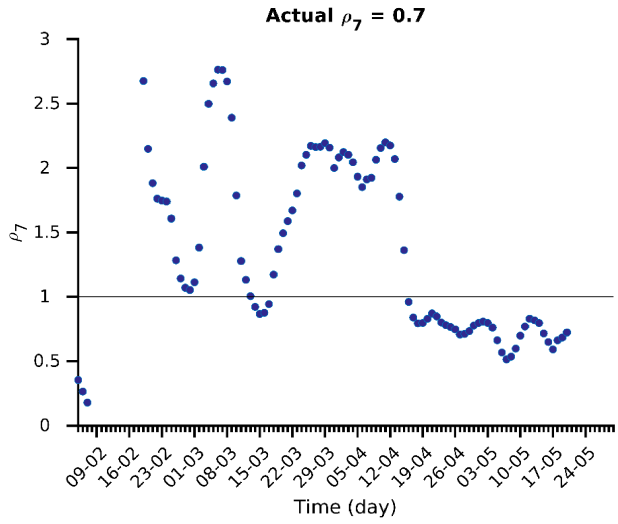


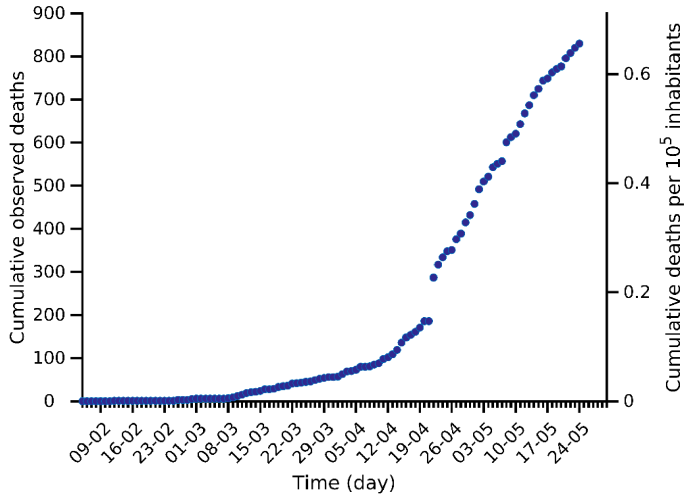


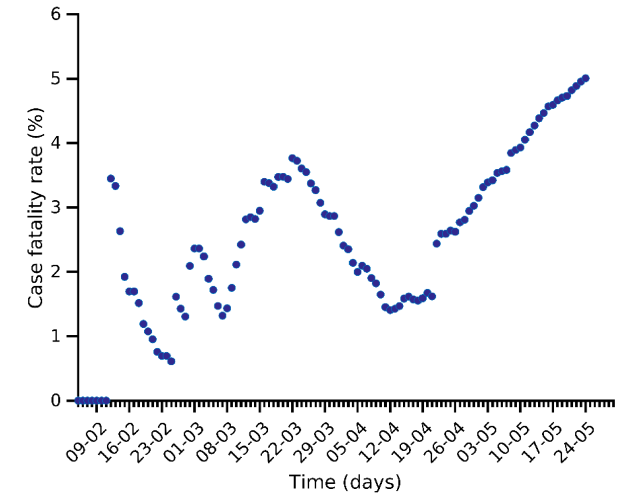




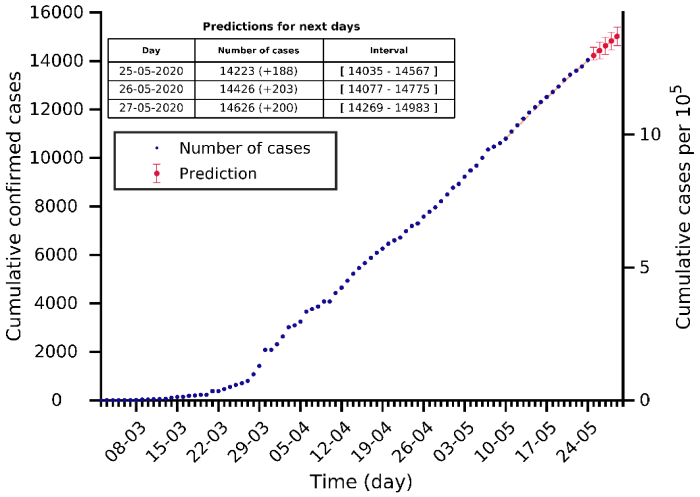


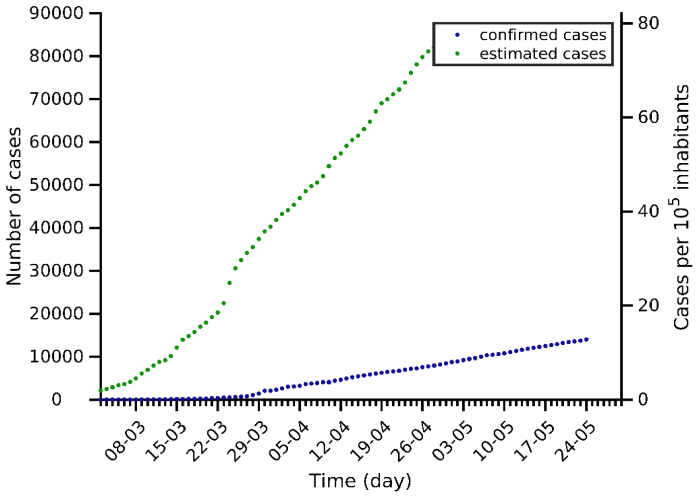


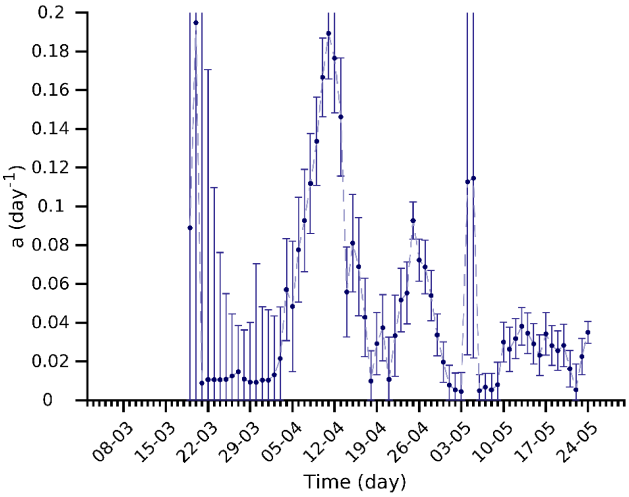


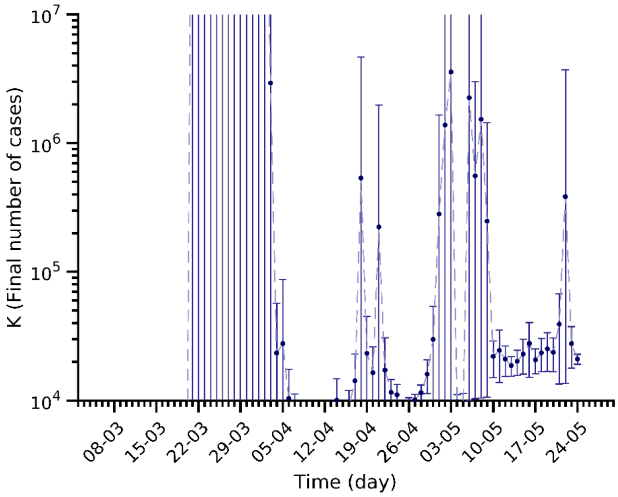


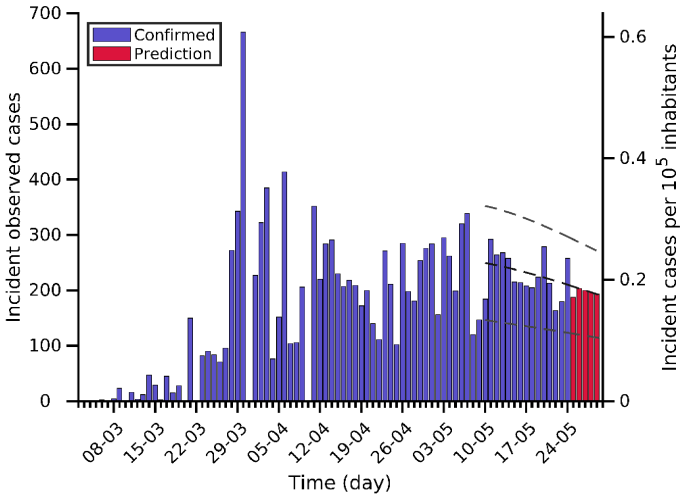
Philippines 24-05-2020. Population: 109.6M. Current cumulated incidence: $13/10^5$

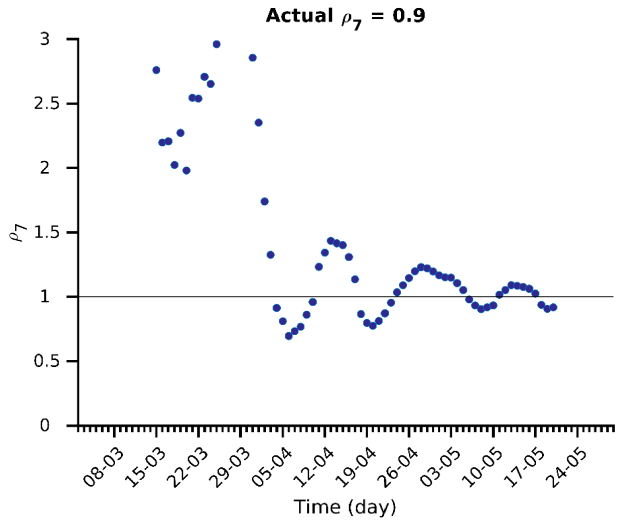


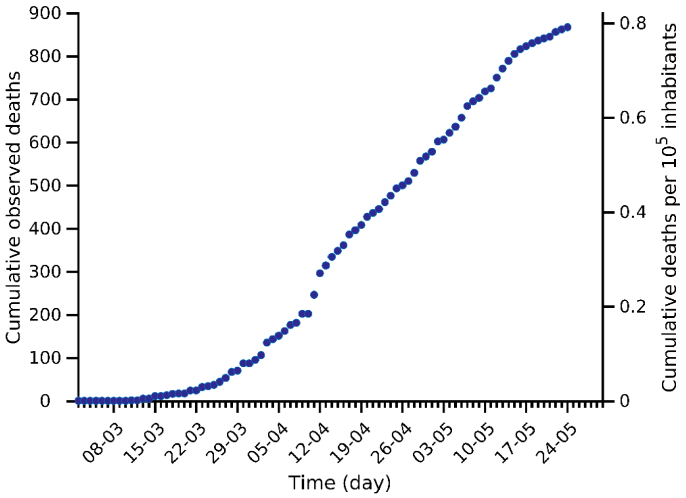


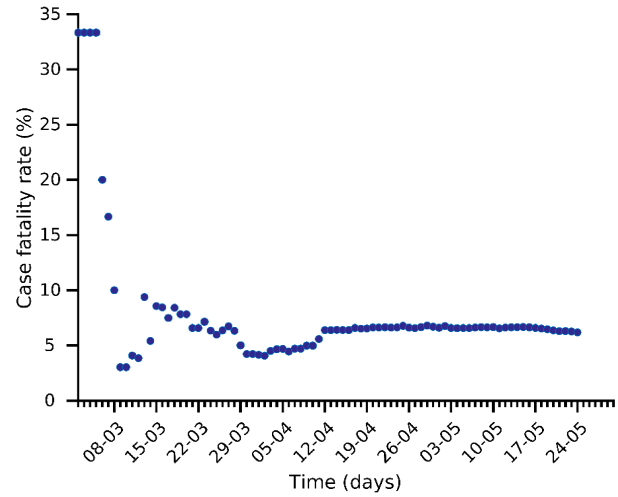








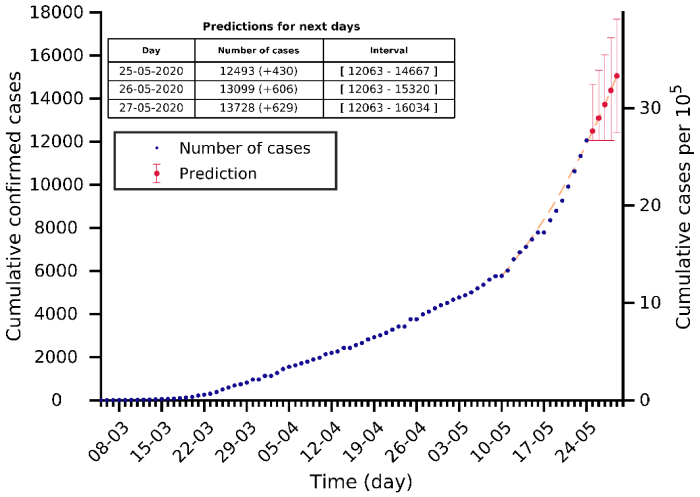


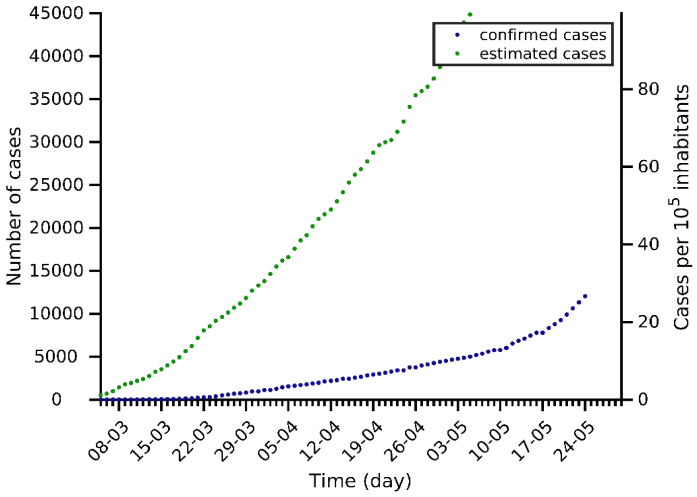


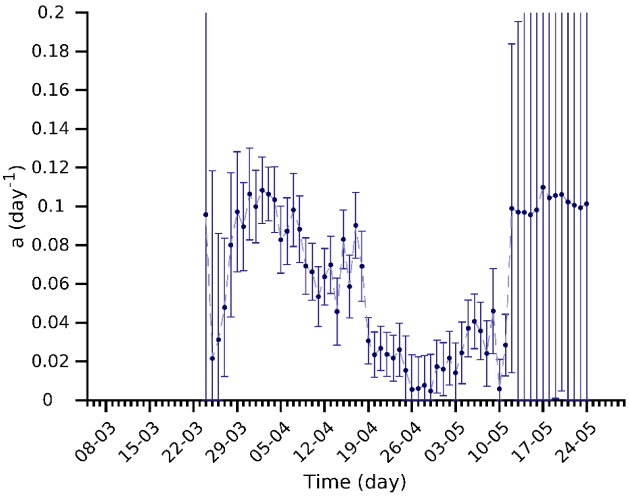
Argentina 24-05-2020. Population: 45.2M. Current cumulated incidence: 27/10^5

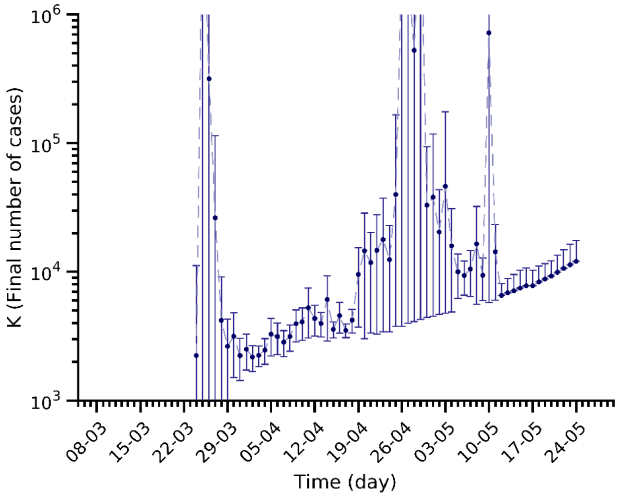
Predictions for next days

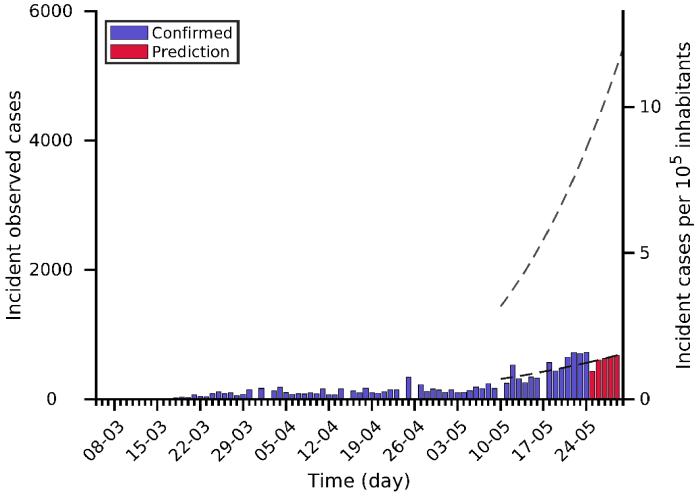
Day	Number of cases	Interval
25-05-2020	12493 (+430)	[12063 - 14667]
26-05-2020	13099 (+606)	[12063 - 15320]
27-05-2020	13728 (+629)	[12063 - 16034]

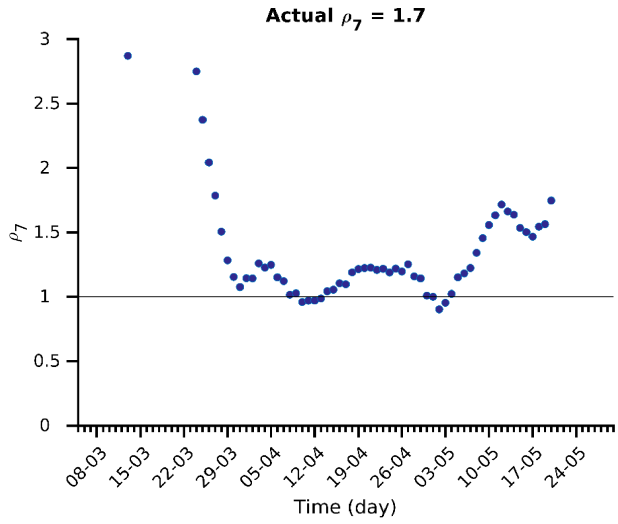


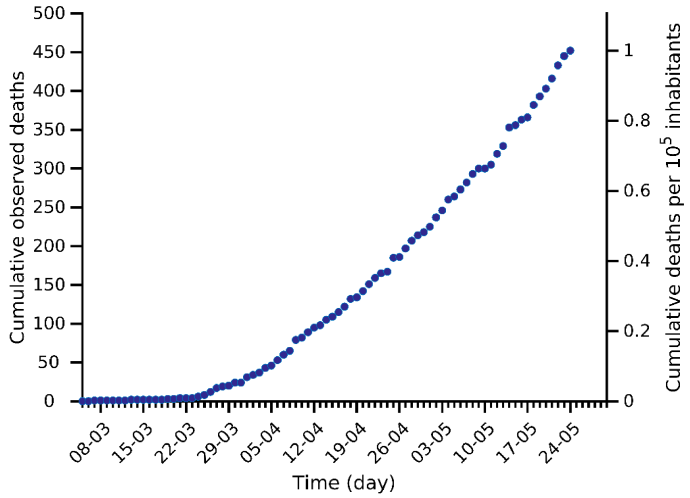


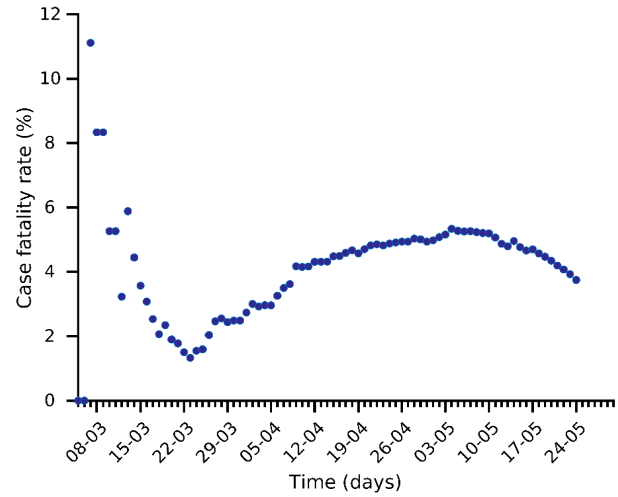




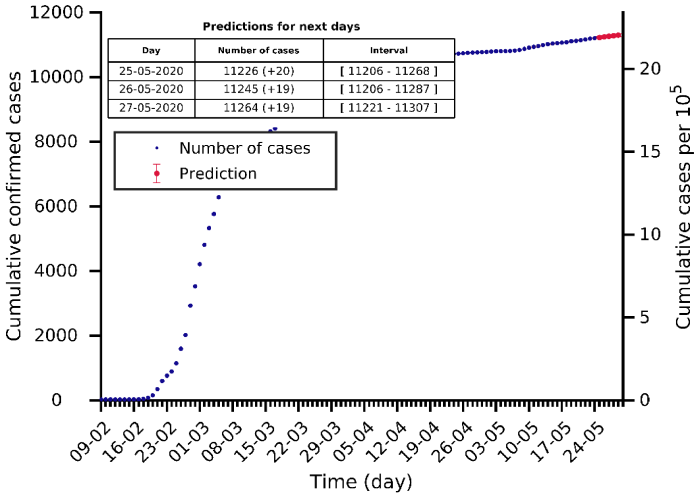


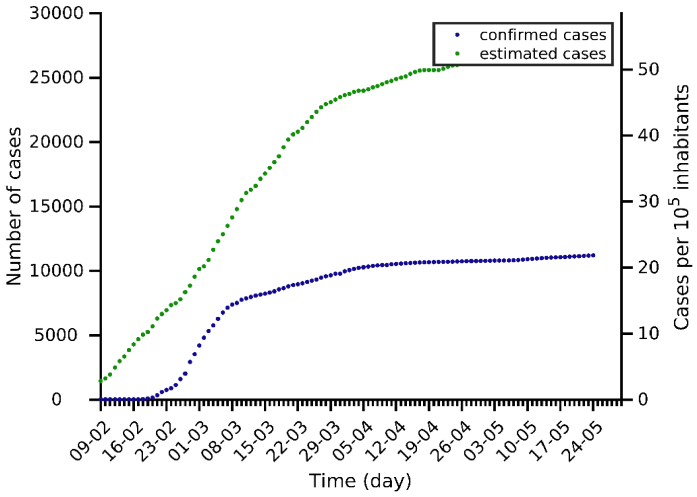


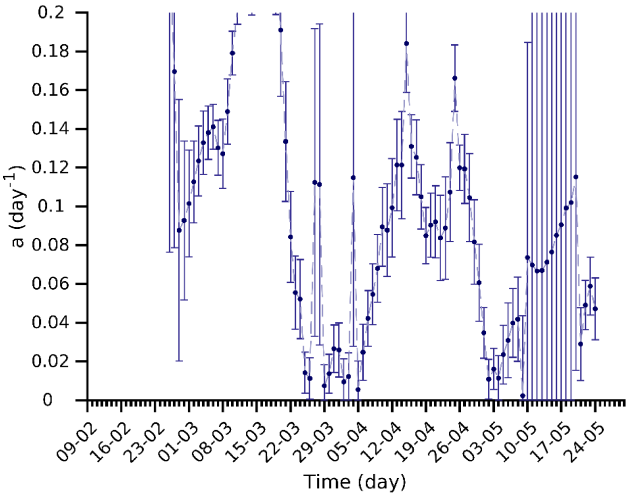


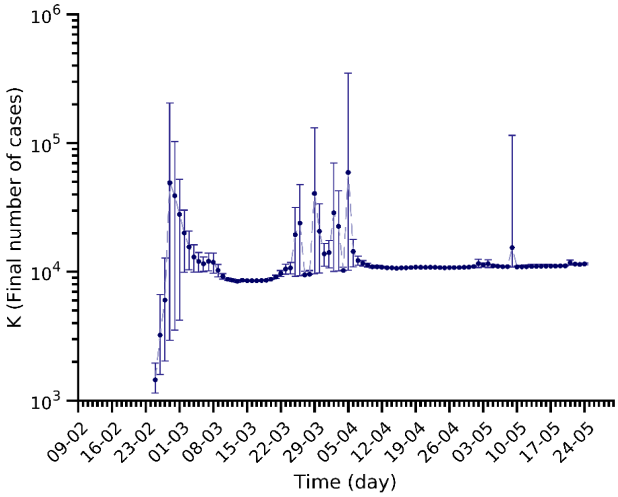


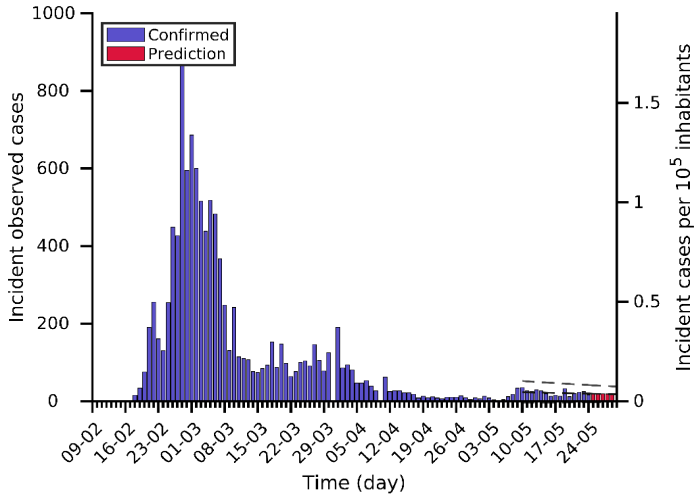
South Korea 24-05-2020. Population: 51.3M. Current cumulated incidence: $22/10^5$

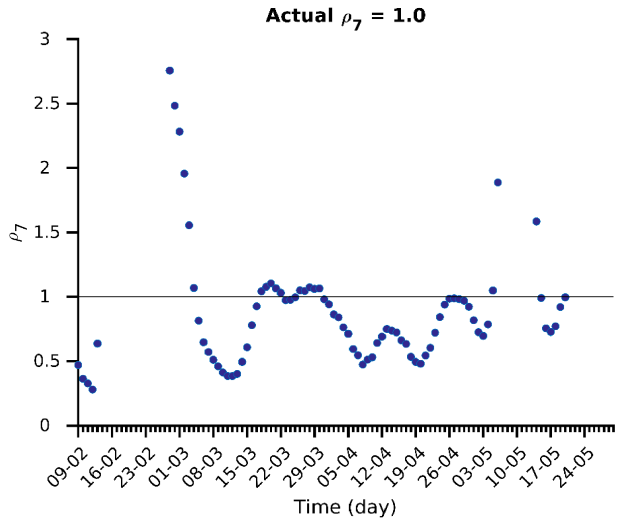


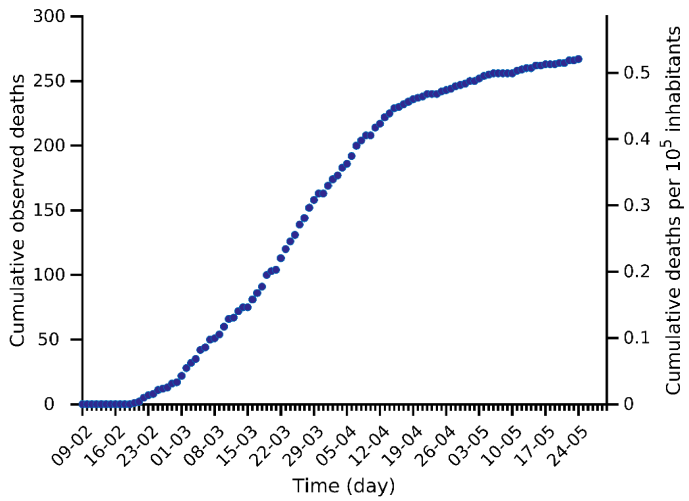


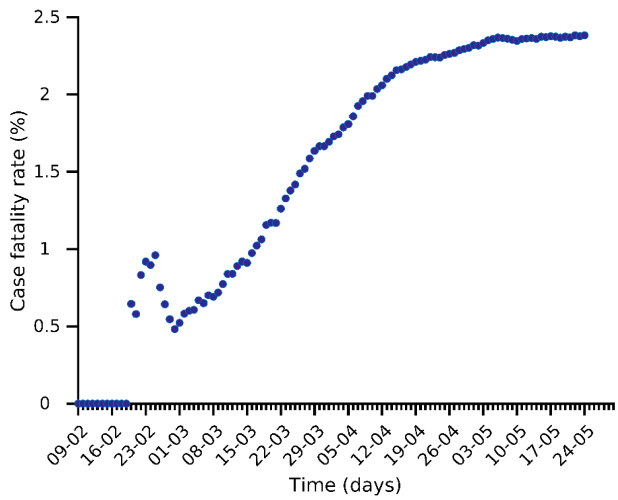




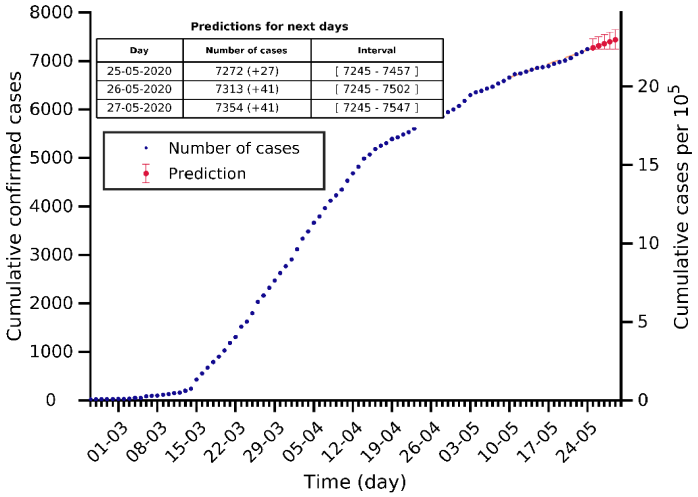


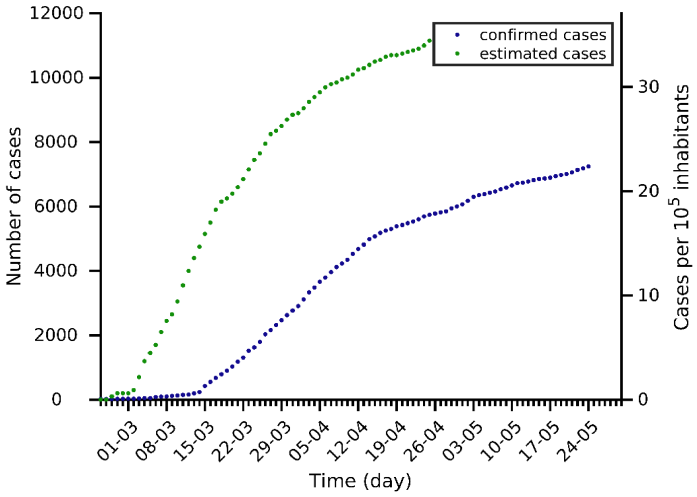


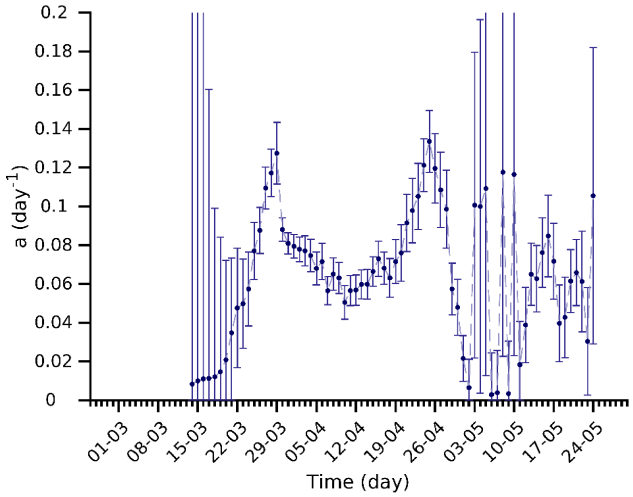


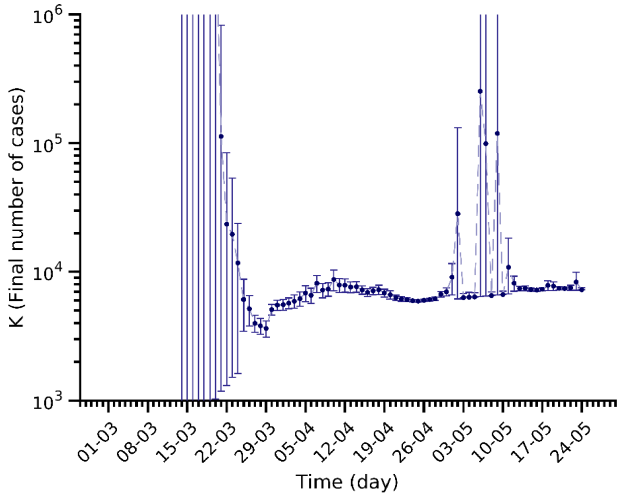


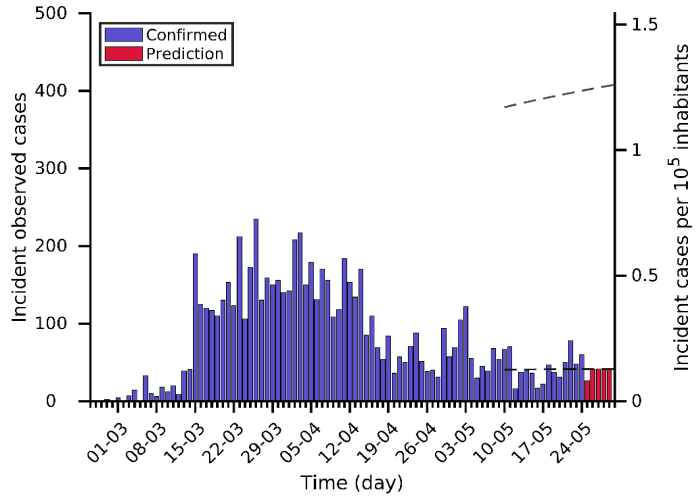
Malaysia 24-05-2020. Population: 32.4M. Current cumulated incidence: $22/10^5$

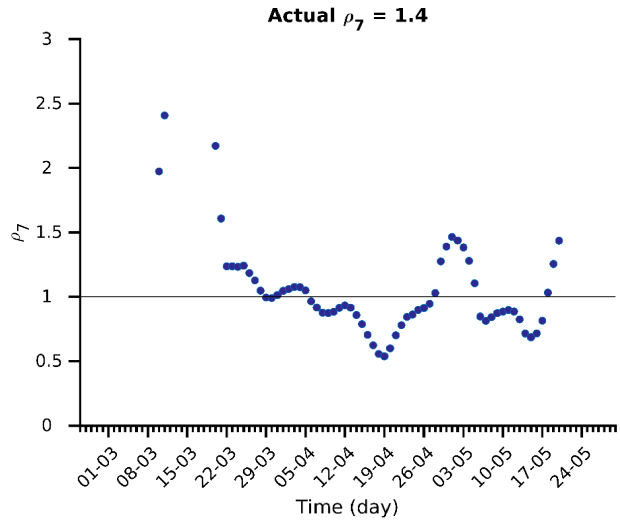


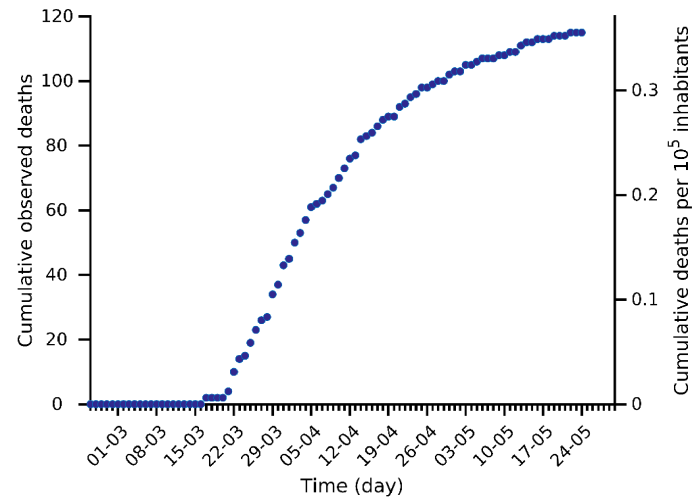


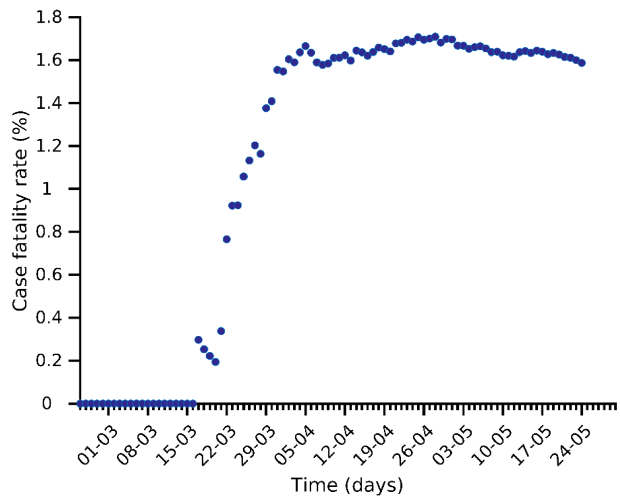




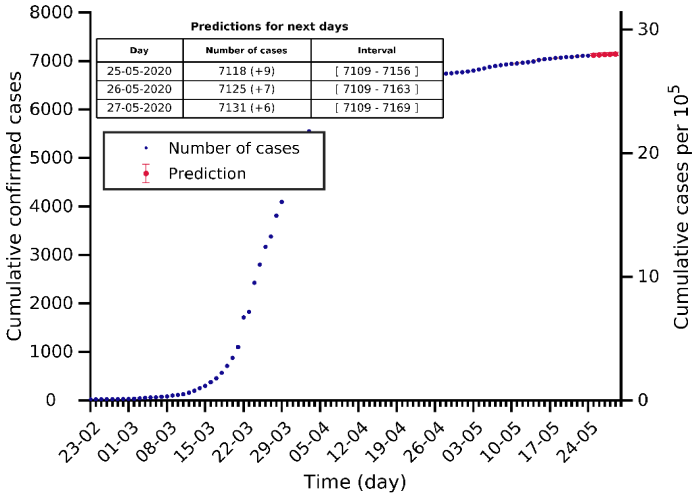


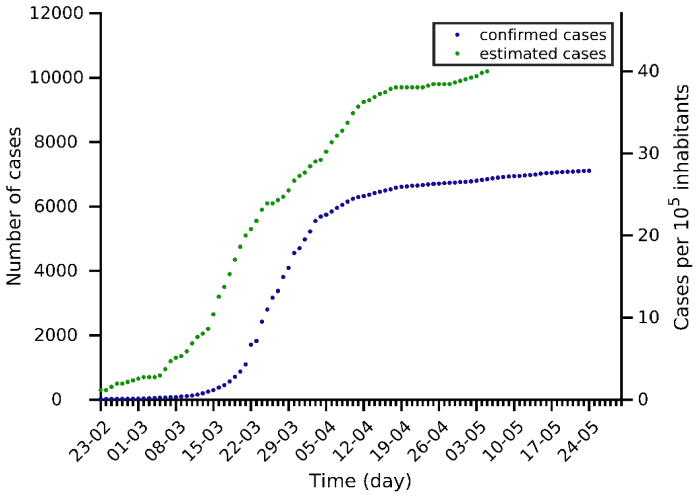


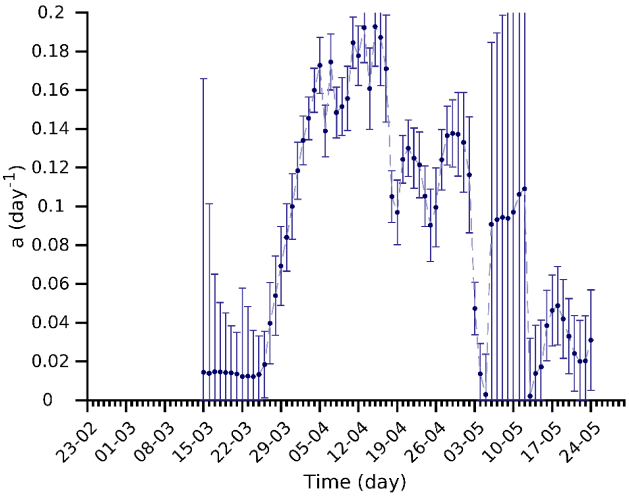


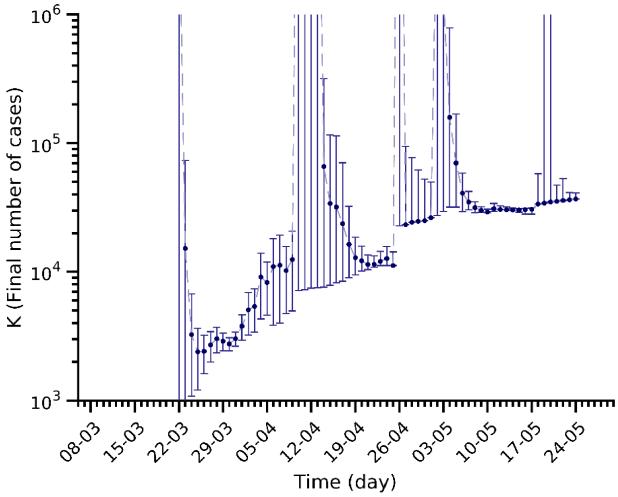


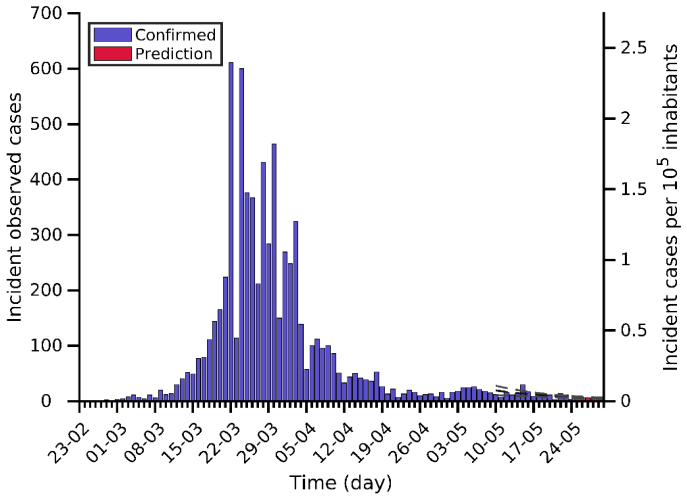
Australia 24-05-2020. Population: 25.5M. Current cumulated incidence: 28/10^5

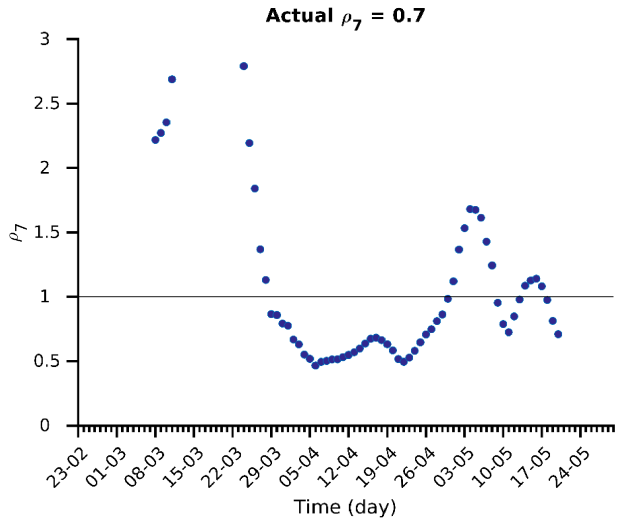


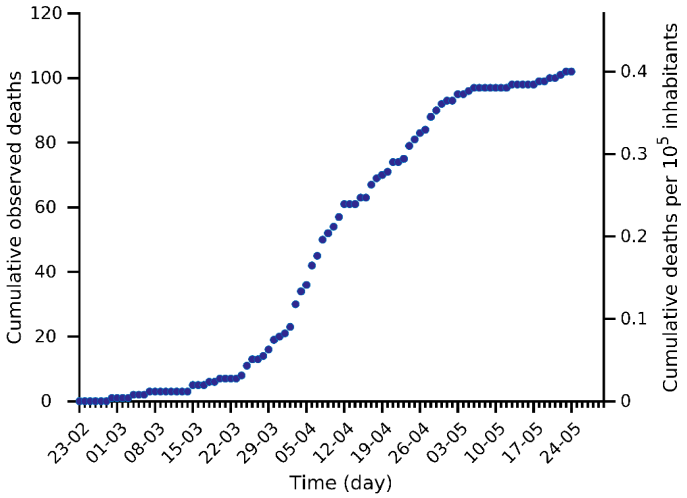


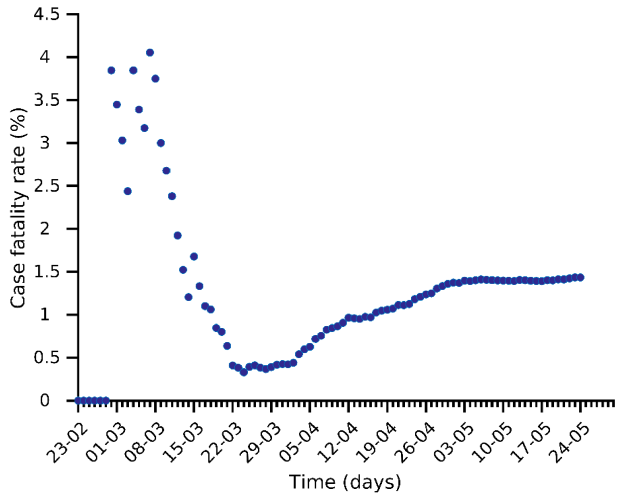




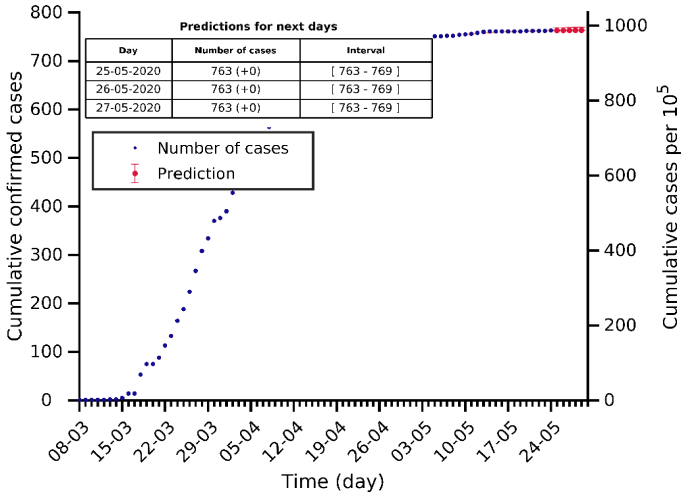


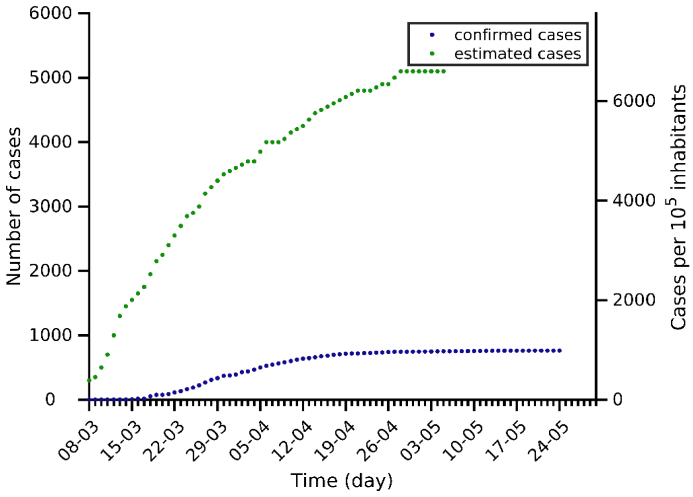


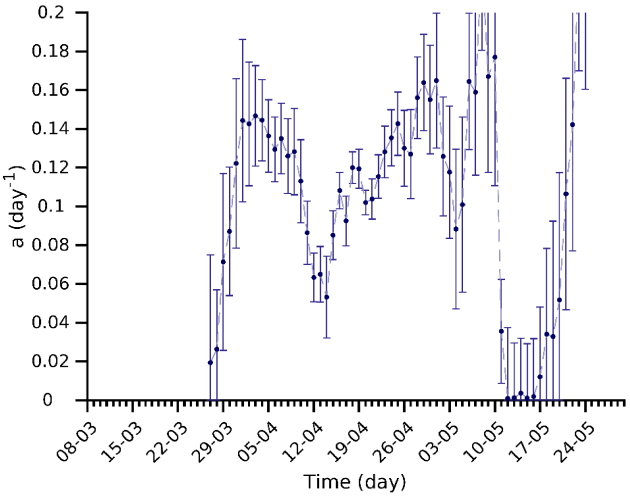


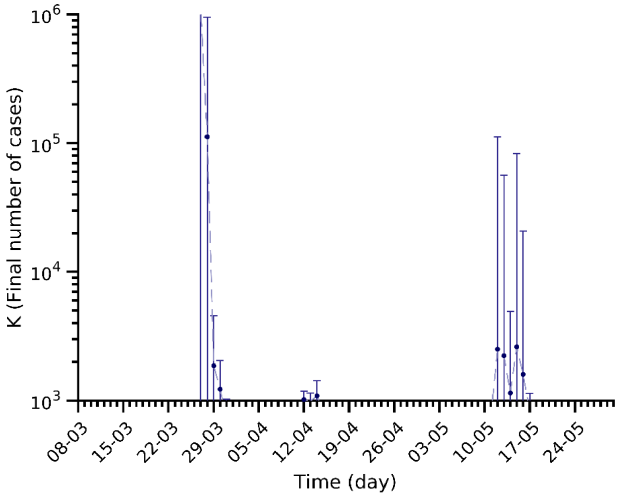


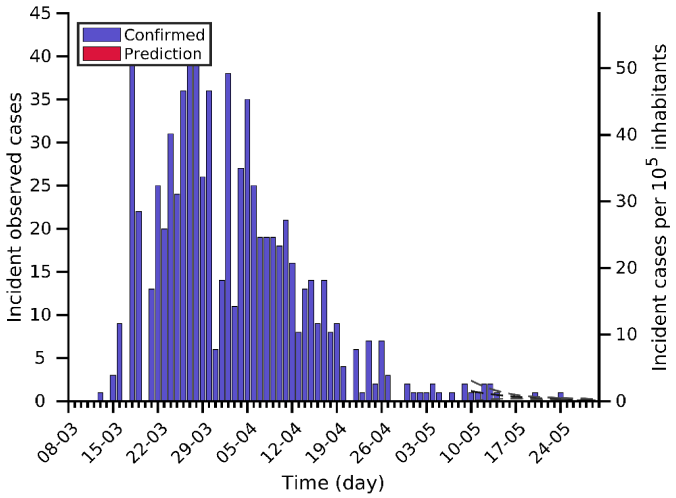
Andorra 24-05-2020. Population: 0.1M. Current cumulated incidence: 988/10^5

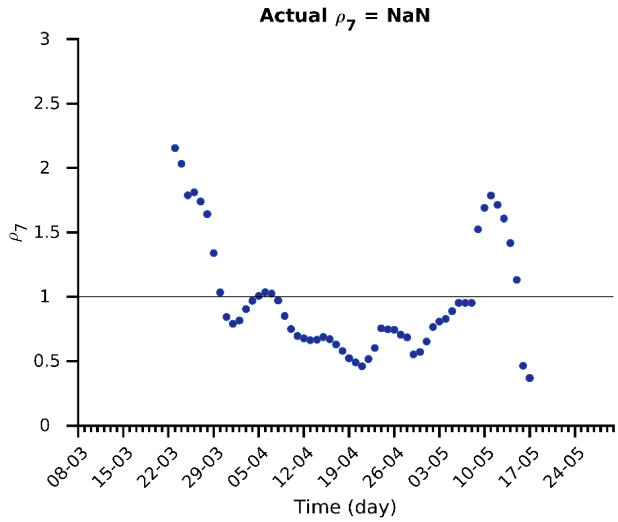


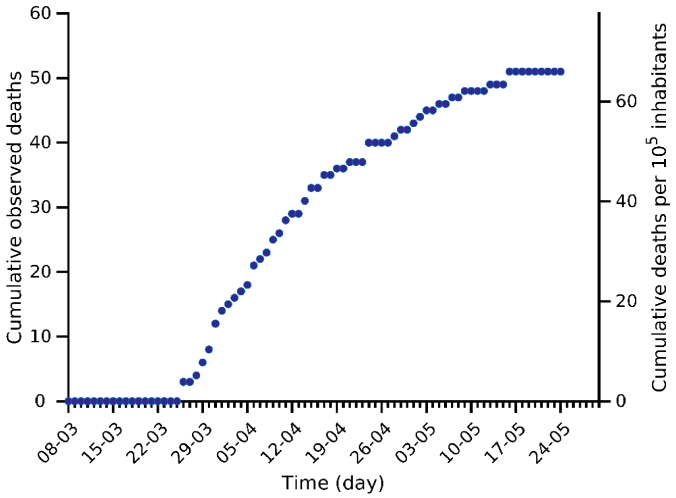


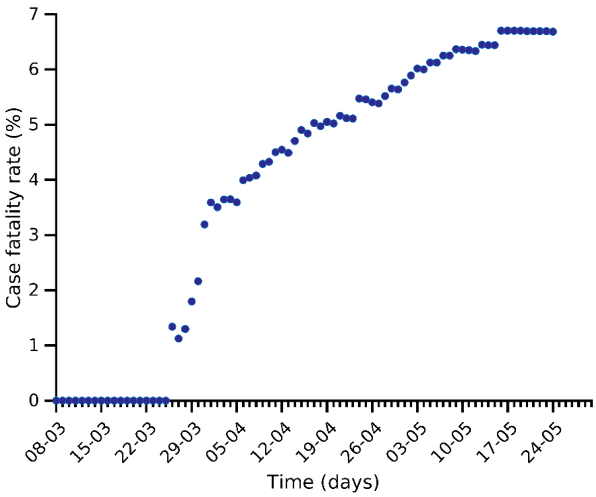








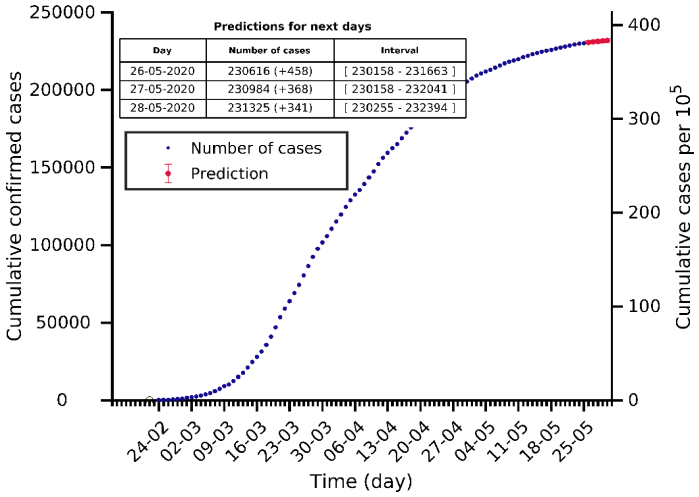


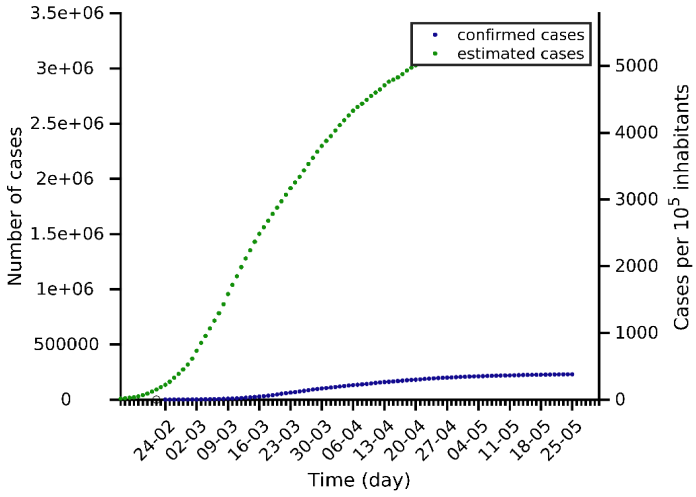


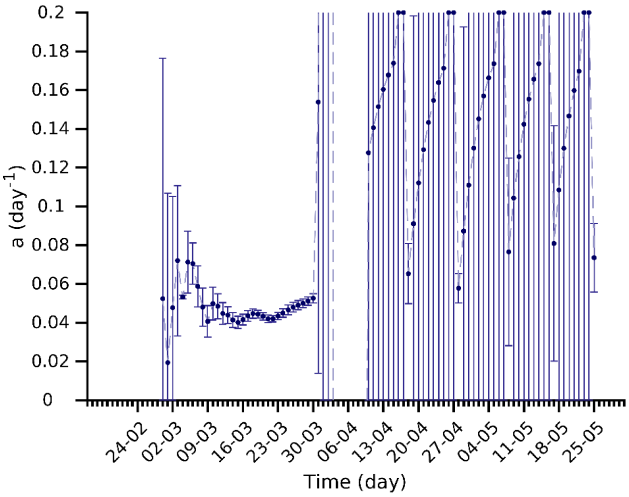
(3) Analysis and prediction of COVID-19 for Italy and its regions

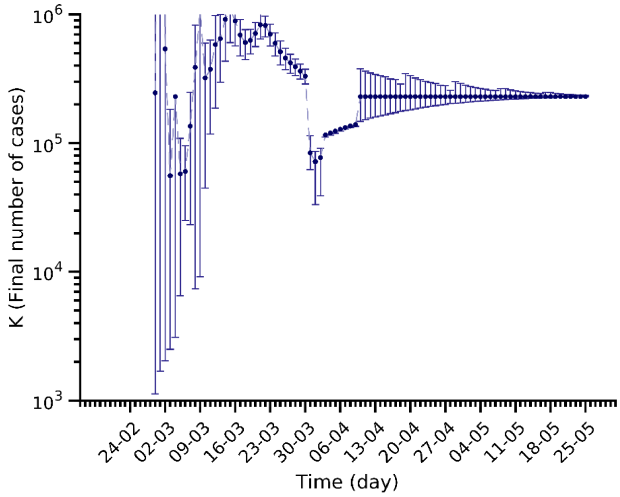
Data obtained from: https://github.com/pcm-dpc/COVID-19/tree/master/dati-andamento-nazionale

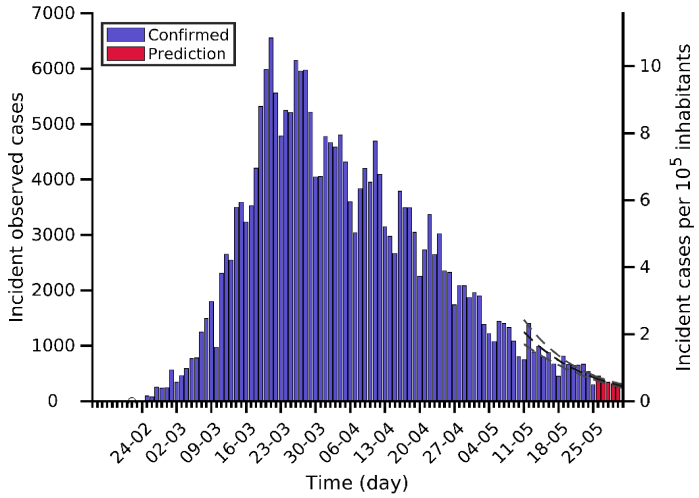
Italy 25-05-2020. Population: 60.5M. Current cumulated incidence: $381/10^5$

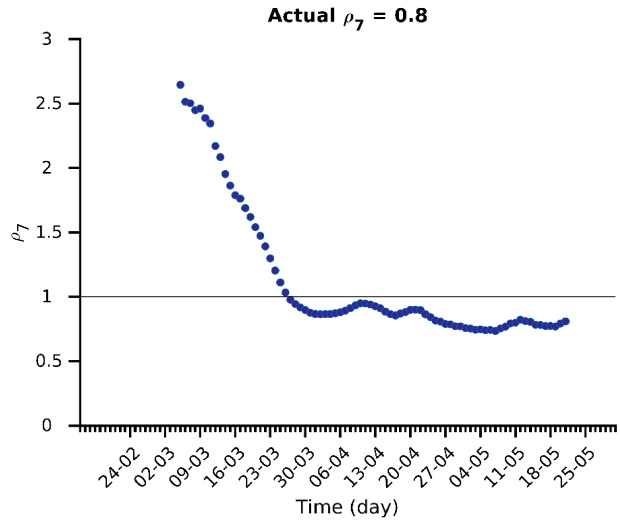


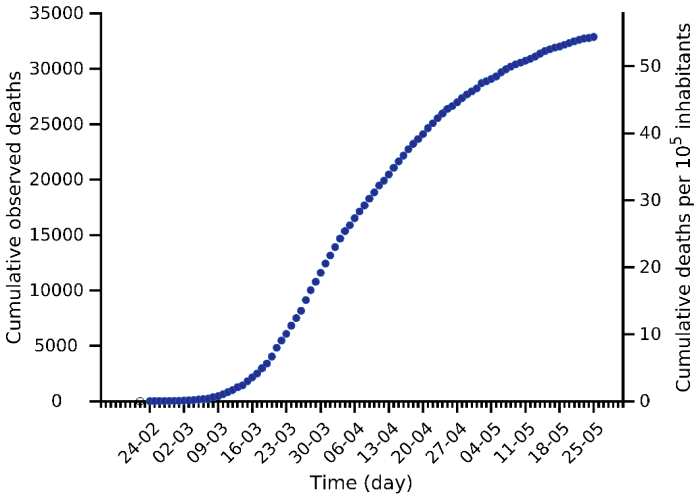


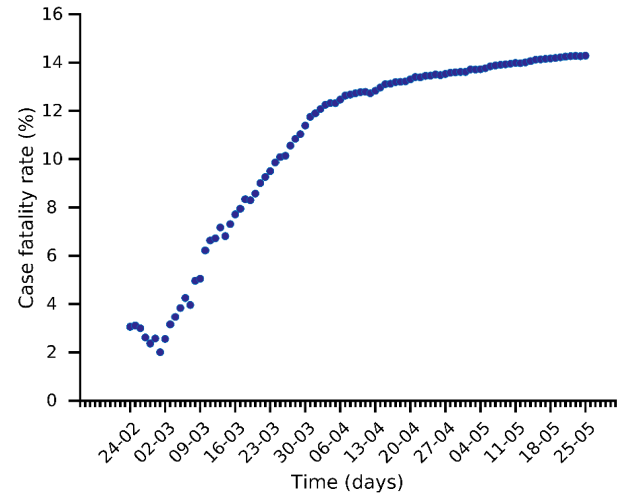




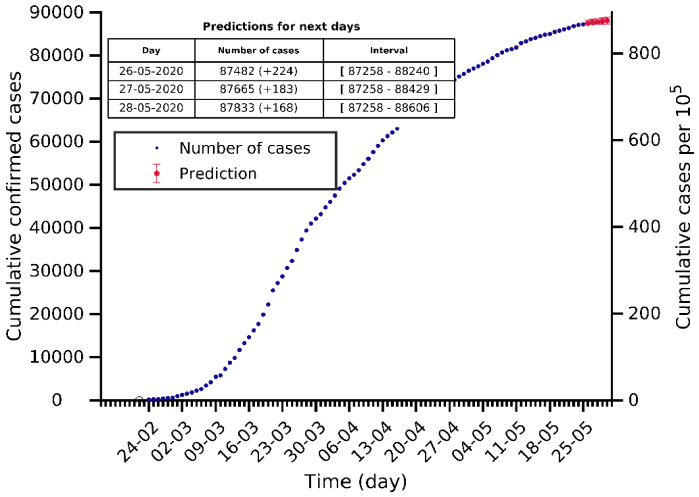


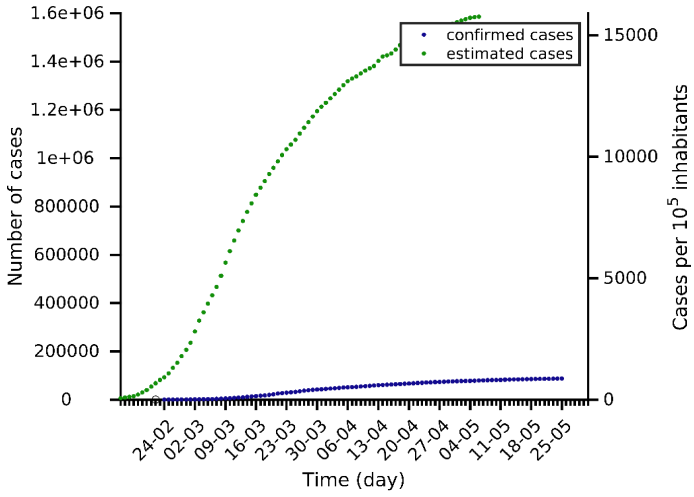


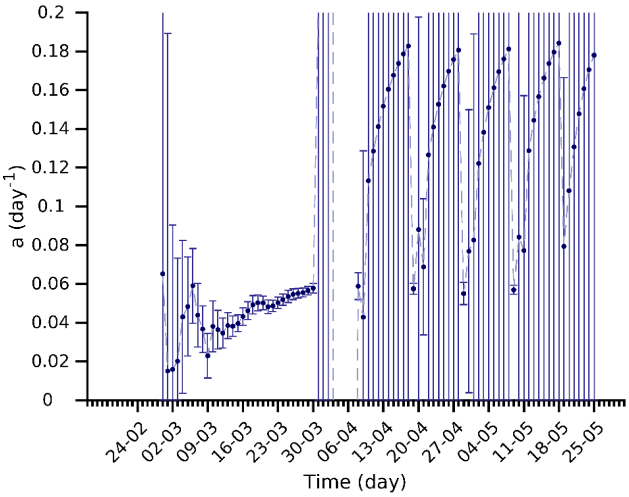


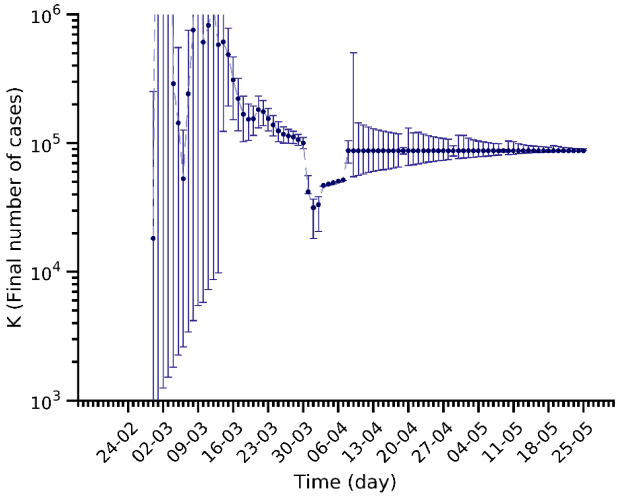


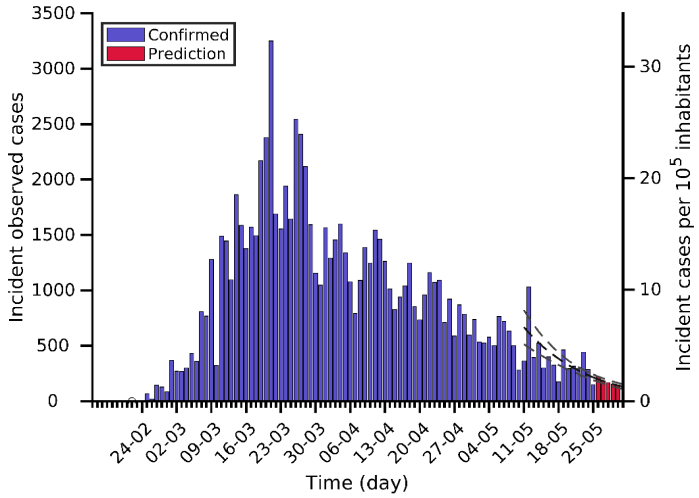
Lombardia 25-05-2020. Population: 10.1M. Current cumulated incidence: 867/10^5

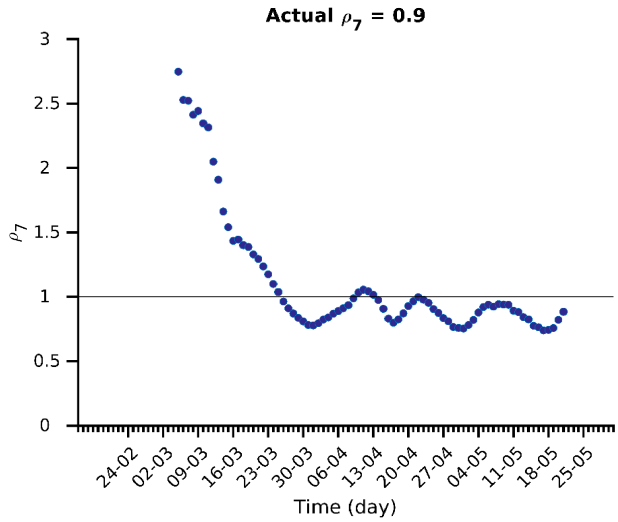


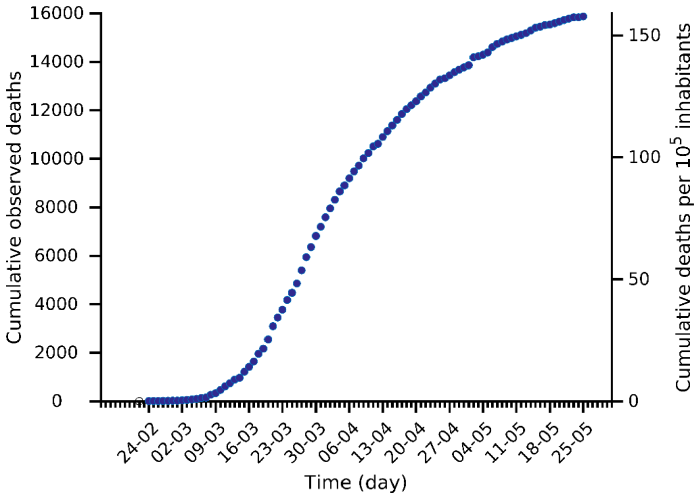


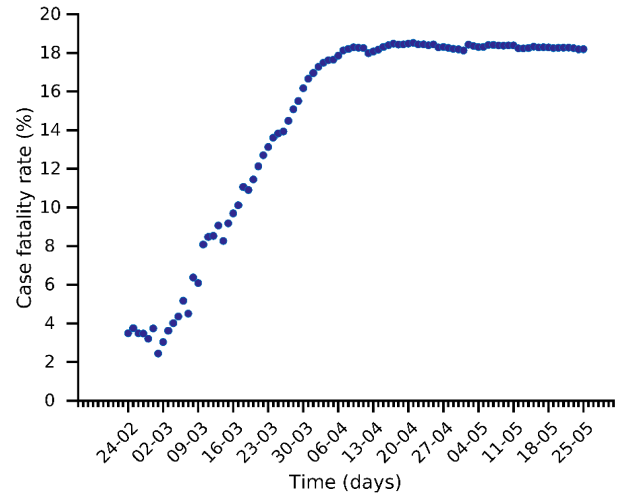




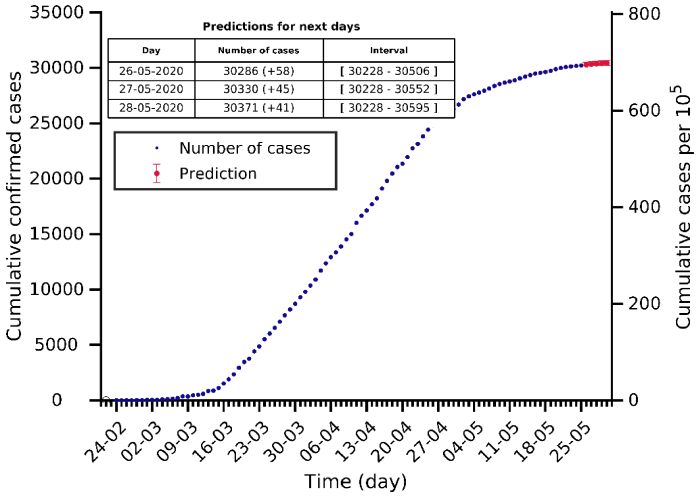


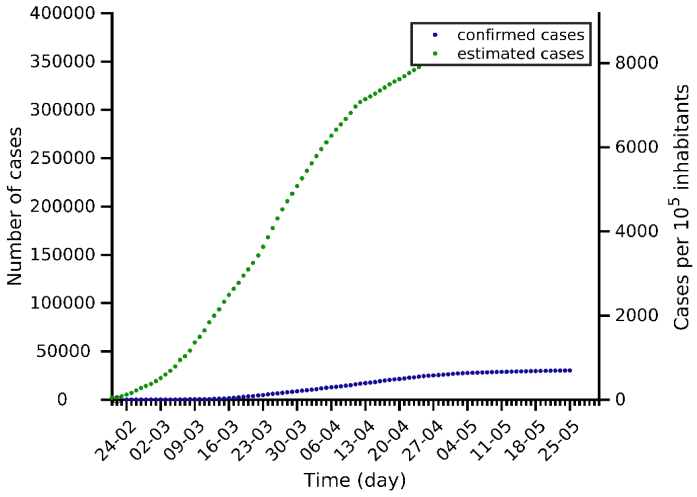


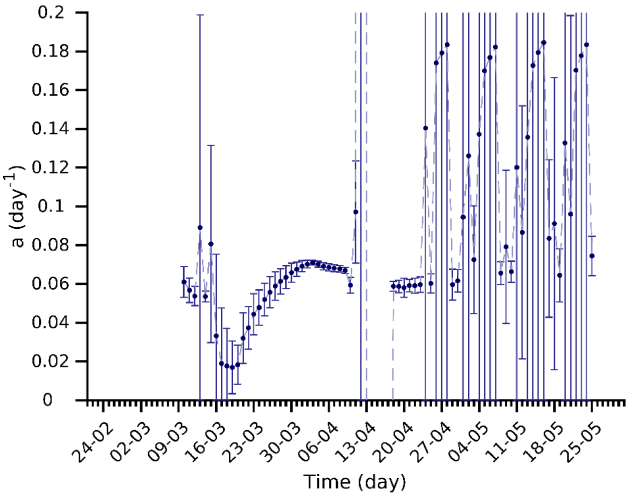


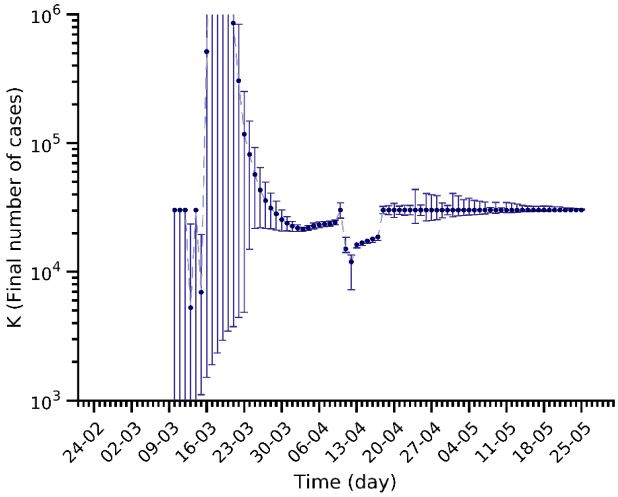


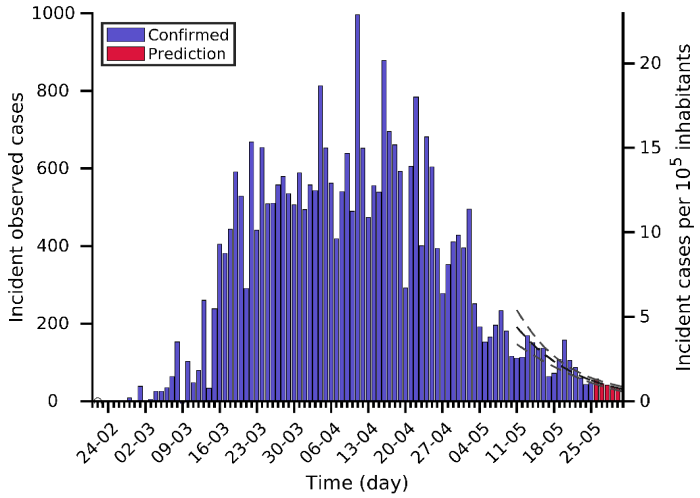
Piemonte 25-05-2020. Population: 4.4M. Current cumulated incidence: 694/10⁵

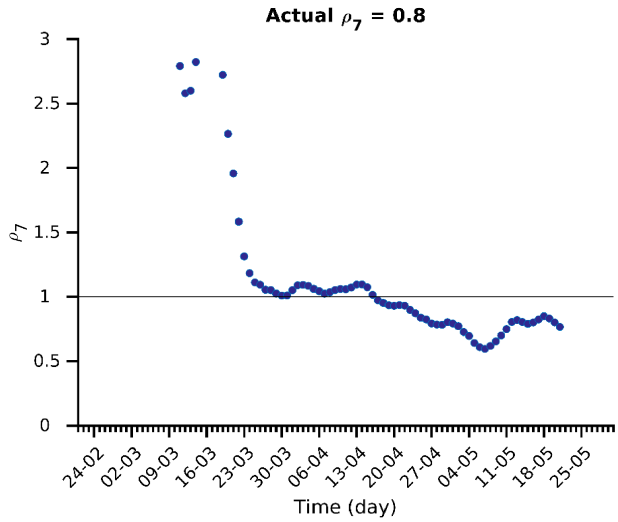


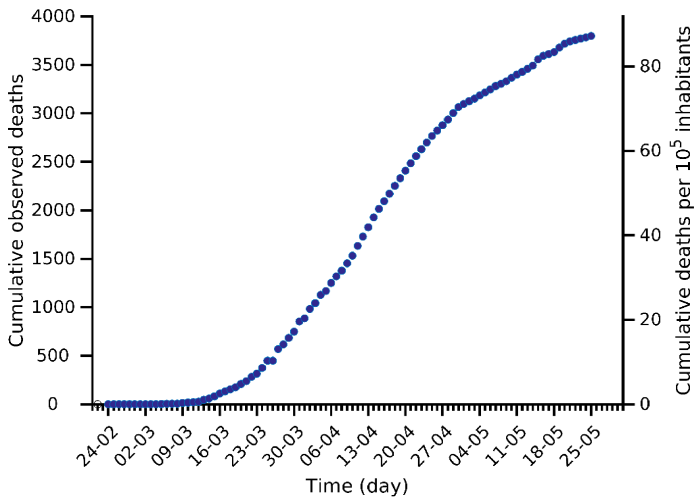


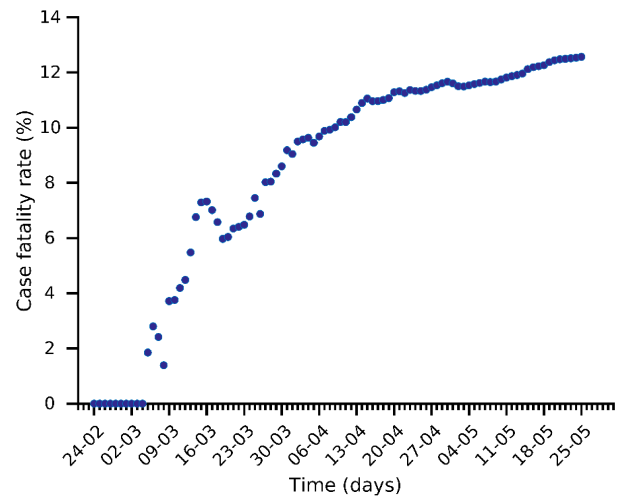




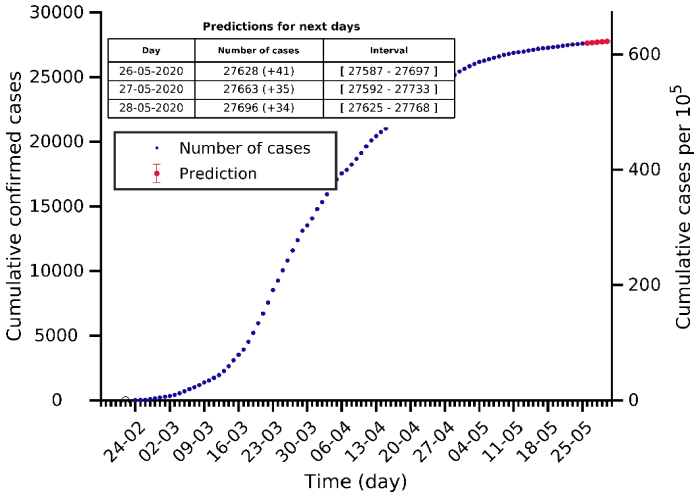


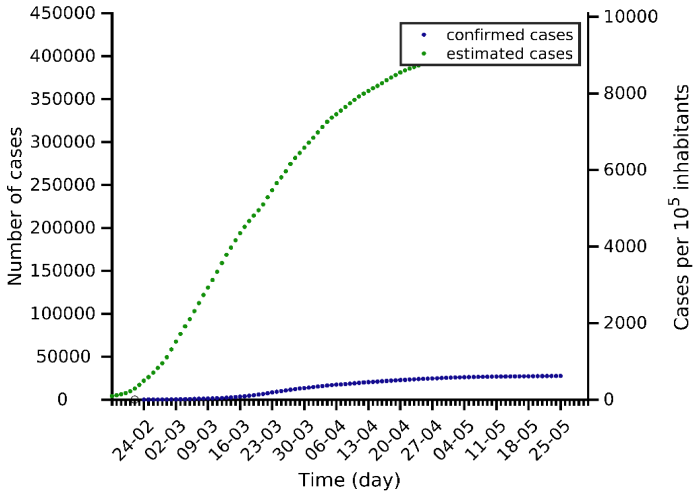


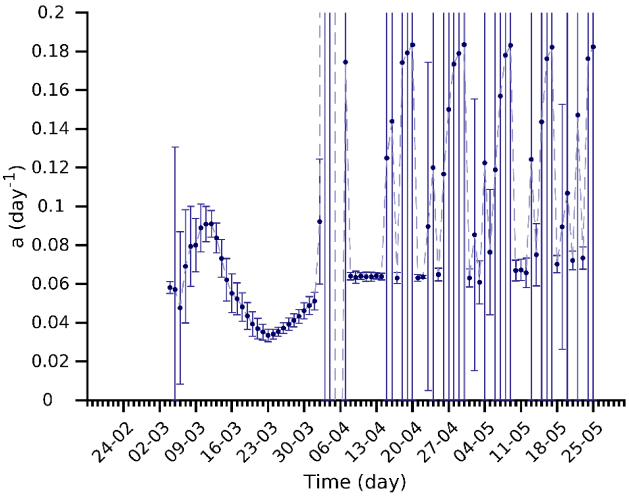


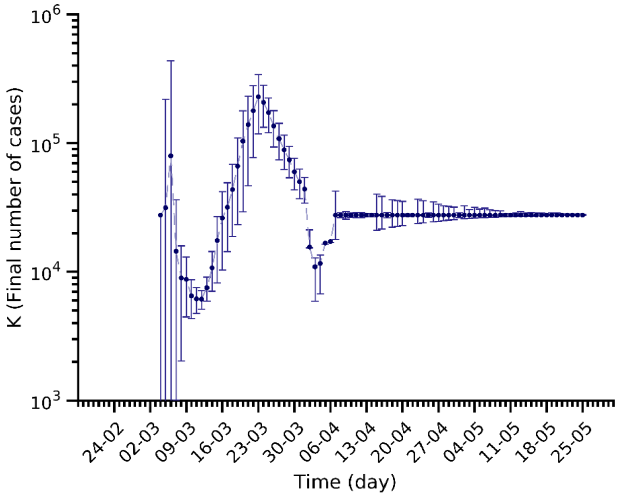


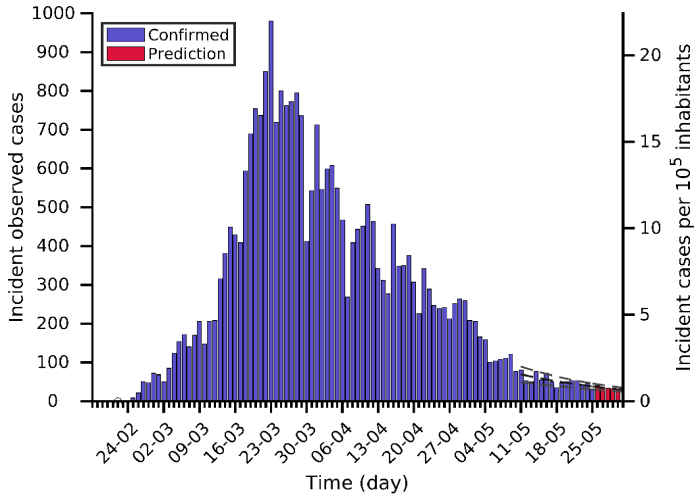
Emilia Romagna 25-05-2020. Population: 4.5M. Current cumulated incidence: 619/1

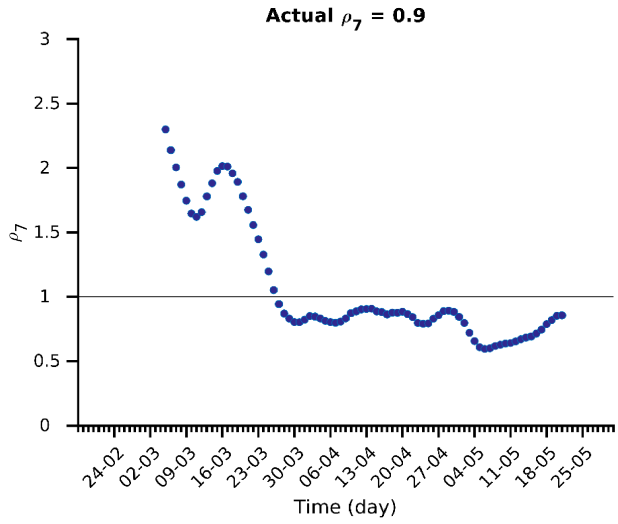


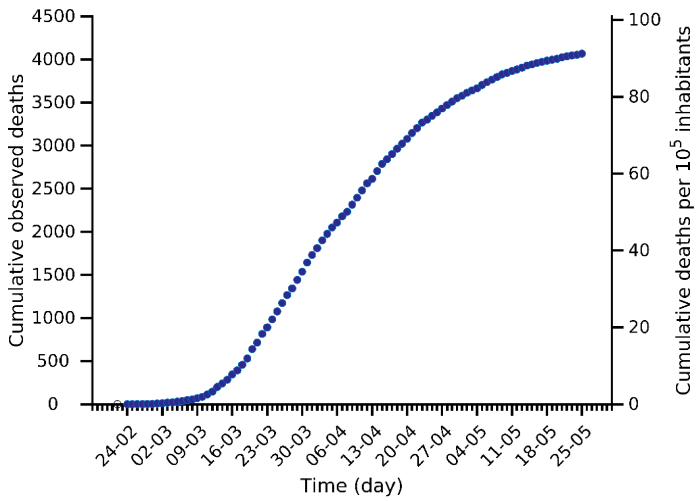


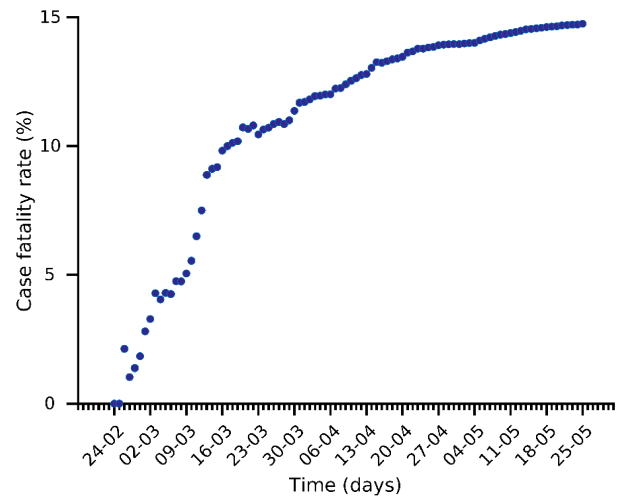




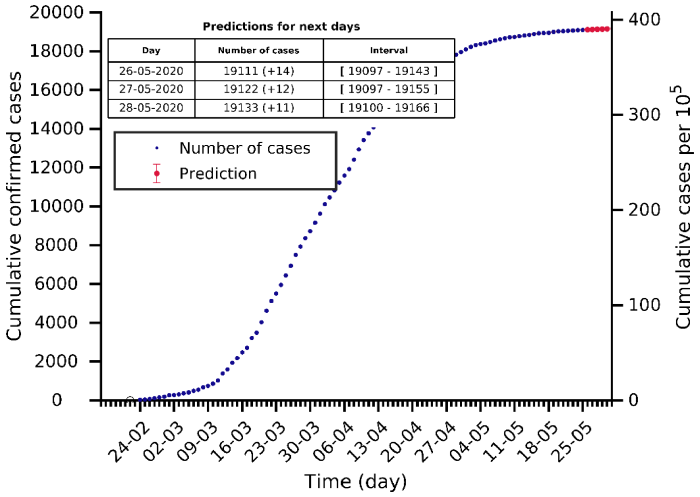


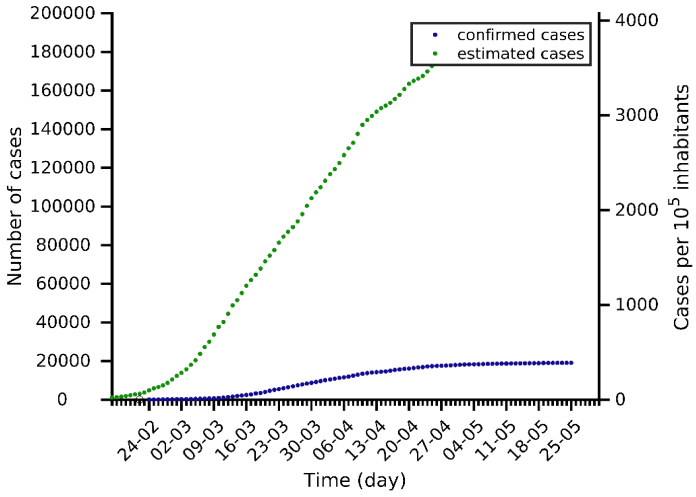


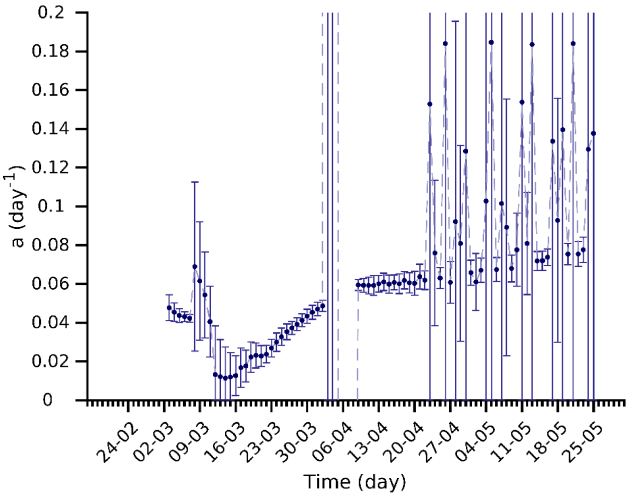


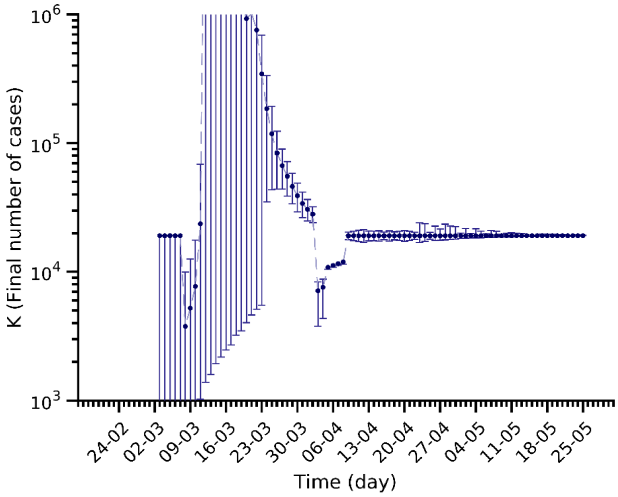


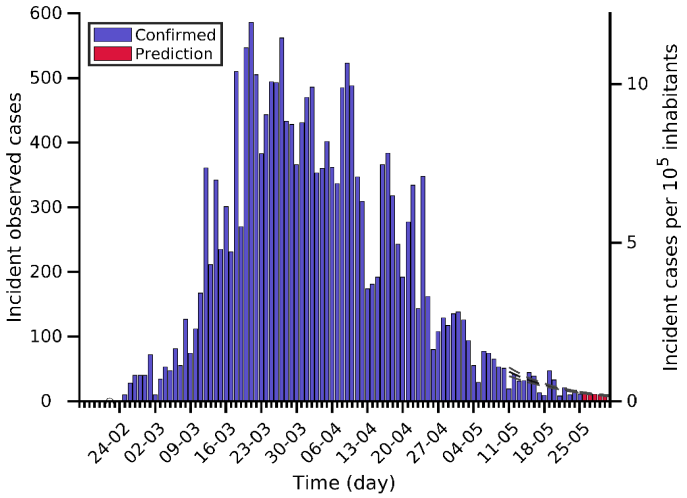
Veneto 25-05-2020. Population: 4.9M. Current cumulated incidence: 389/10⁵

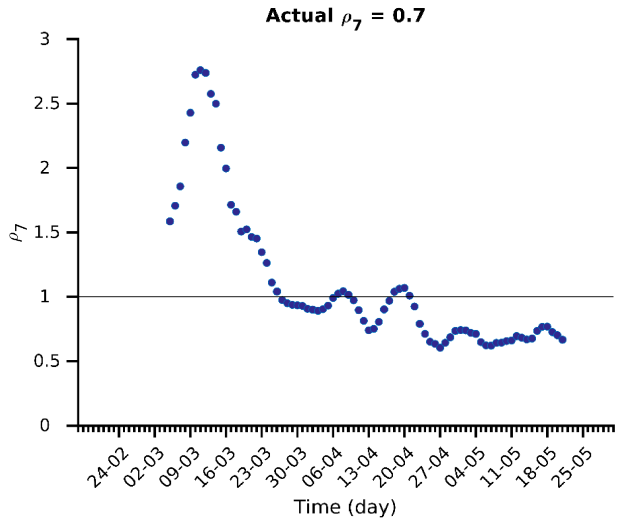


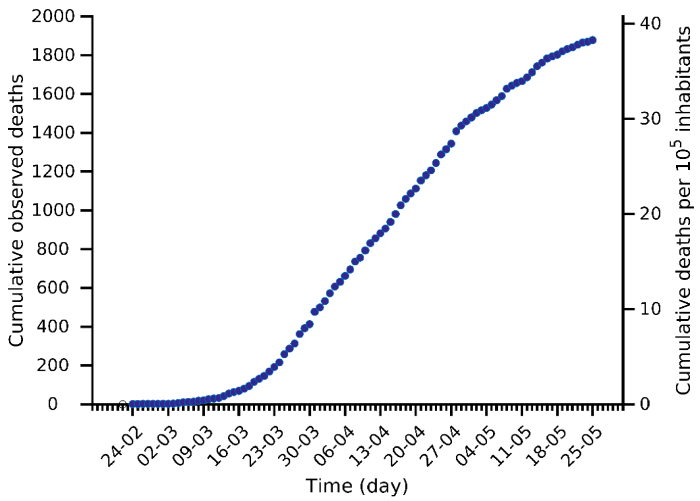


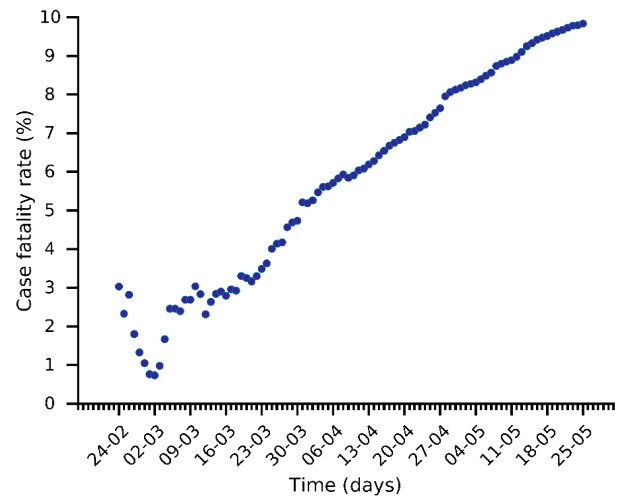




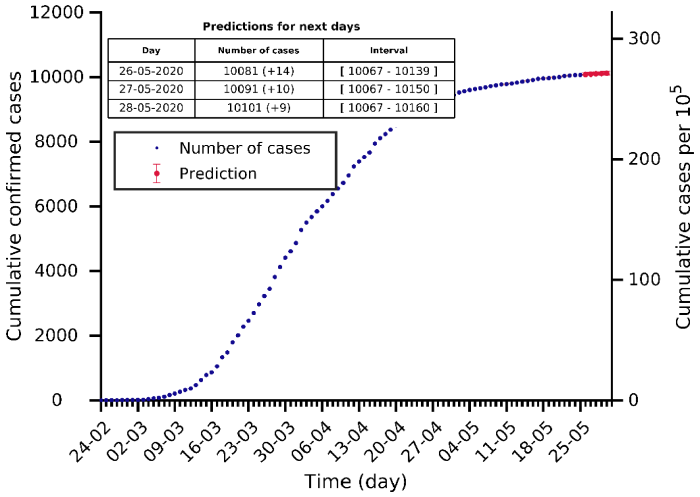


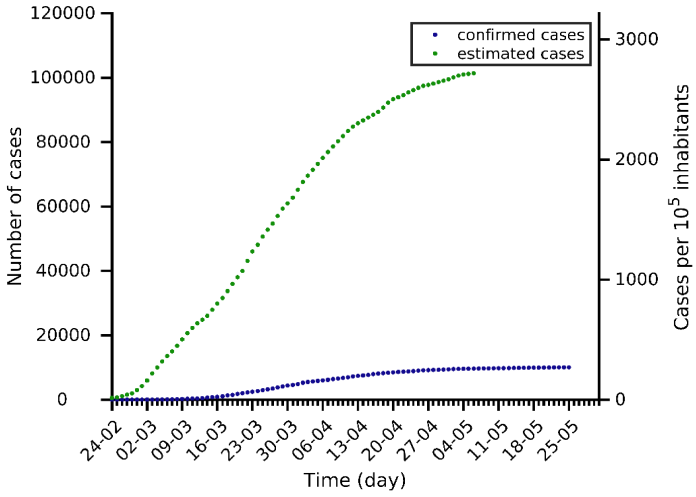


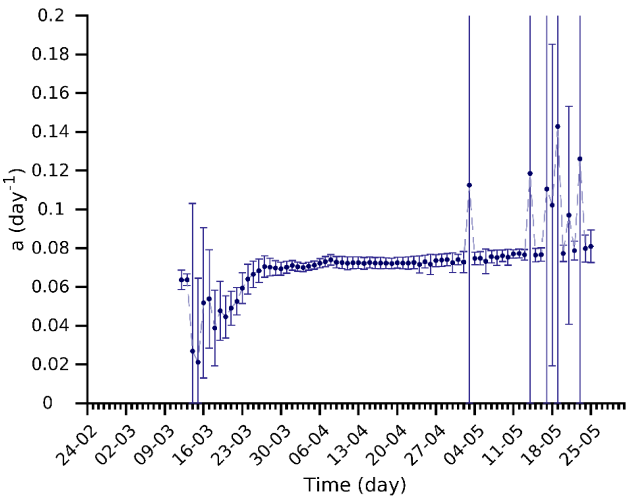


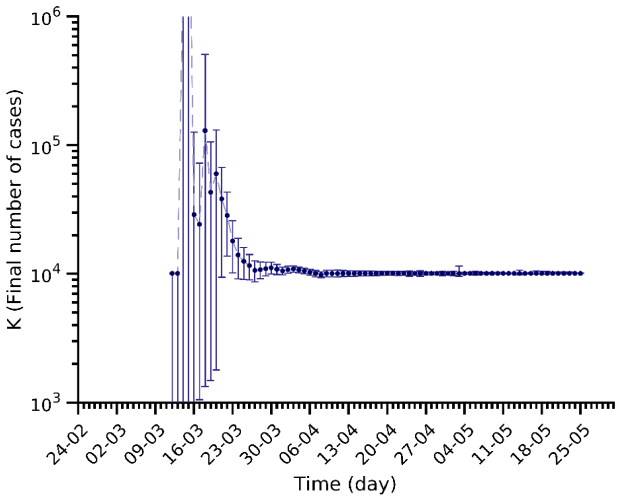


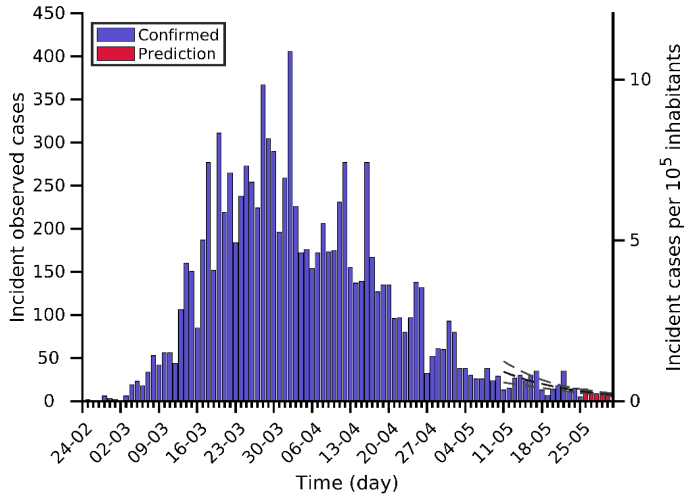
Toscana 25-05-2020. Population: 3.7M. Current cumulated incidence: 270/10⁵

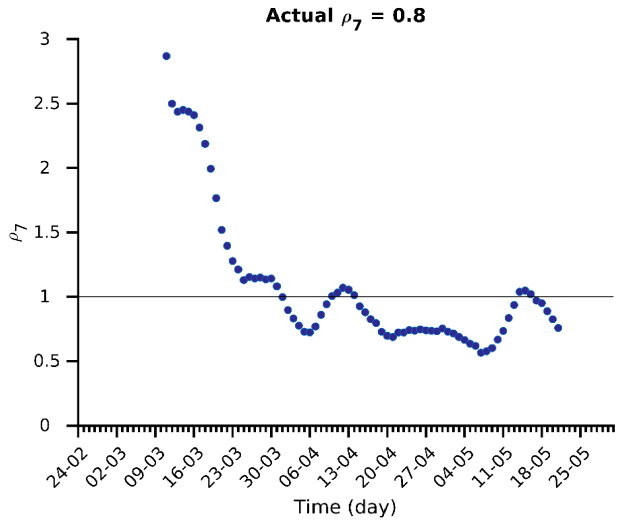


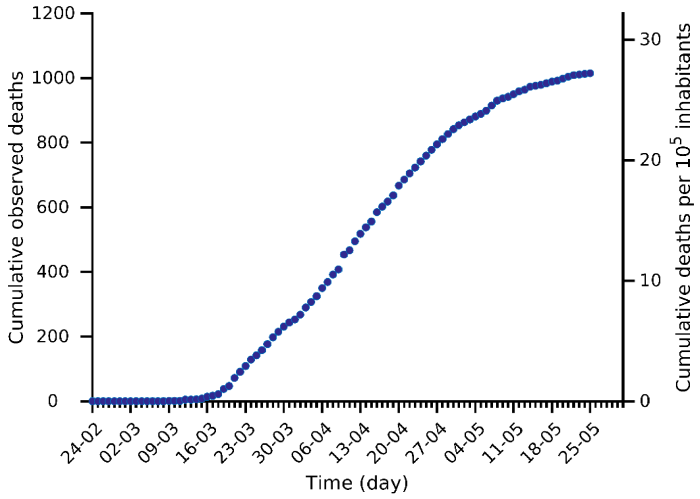


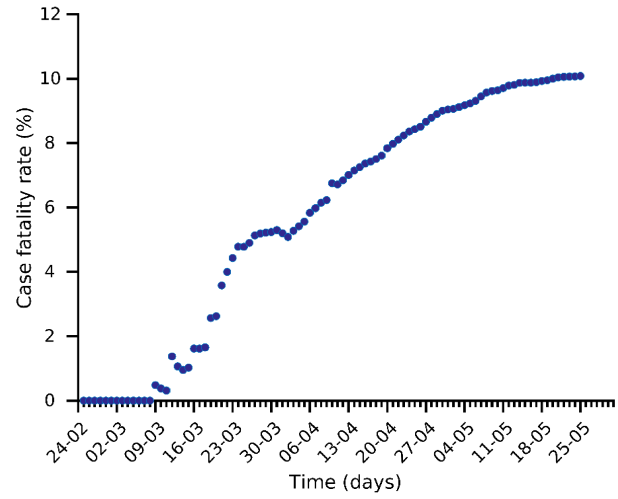




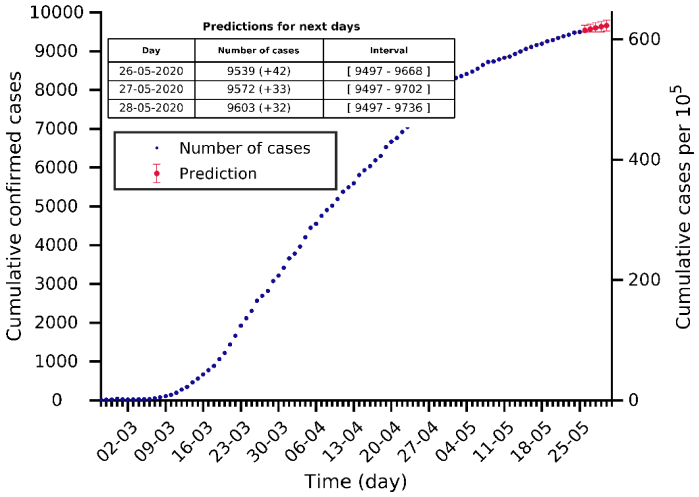


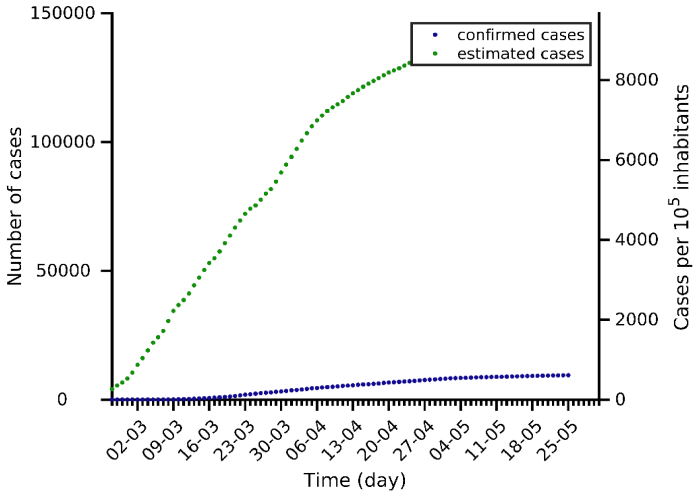


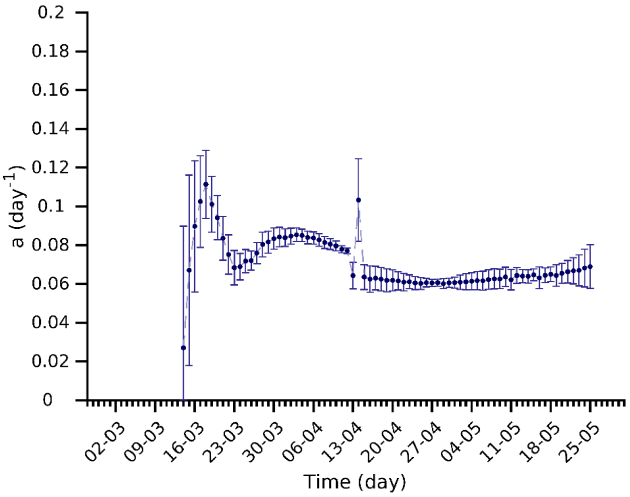


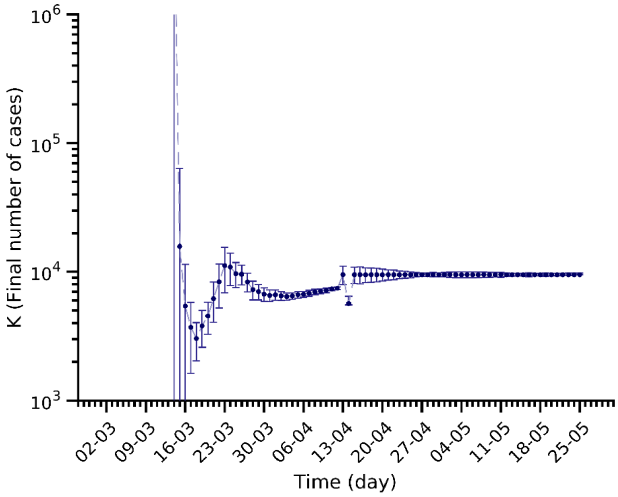


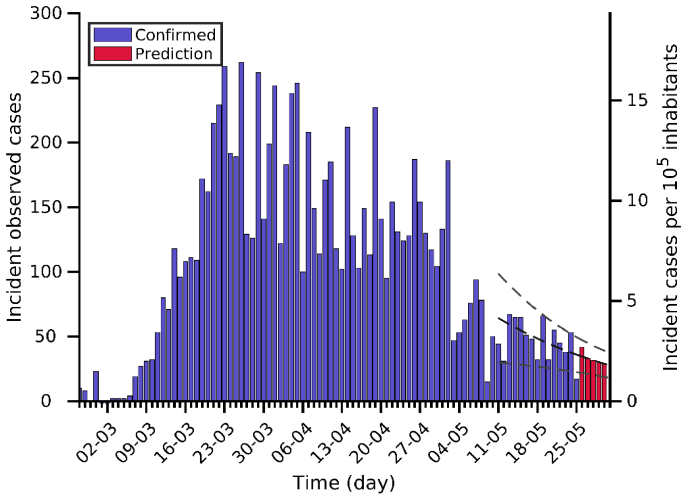
Liguria 25-05-2020. Population: 1.6M. Current cumulated incidence: 612/10⁵

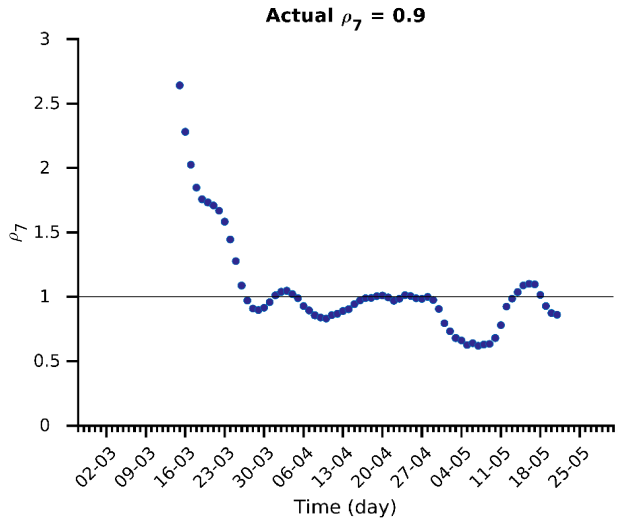


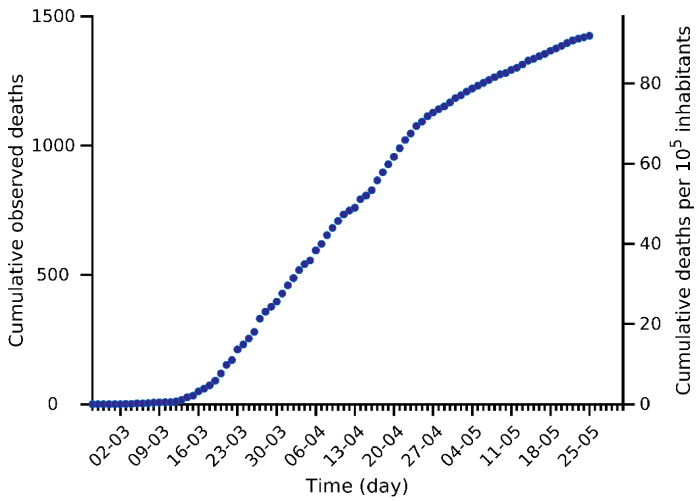


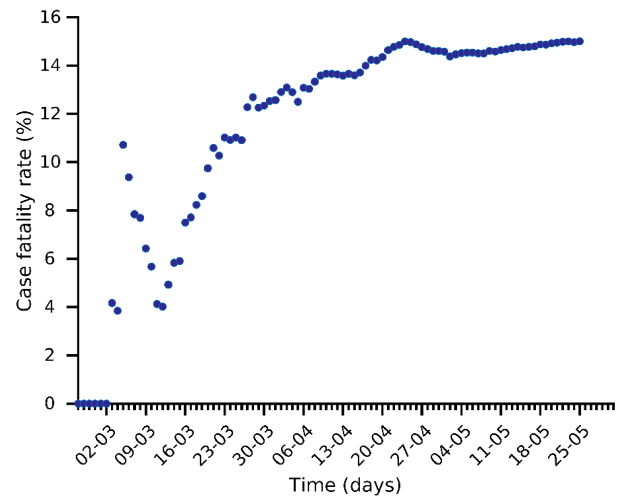




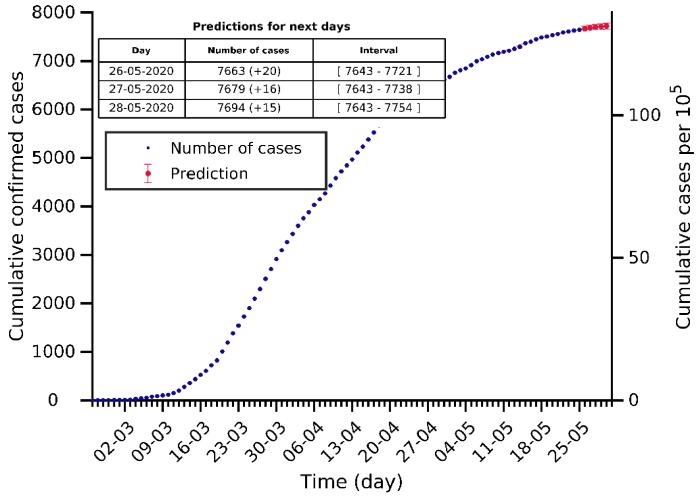


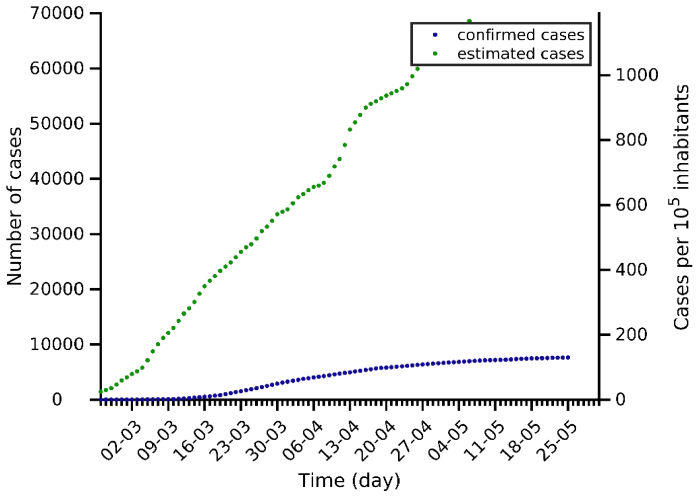


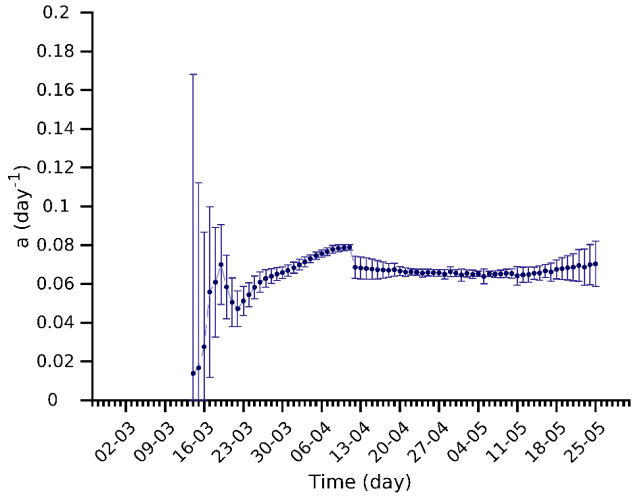


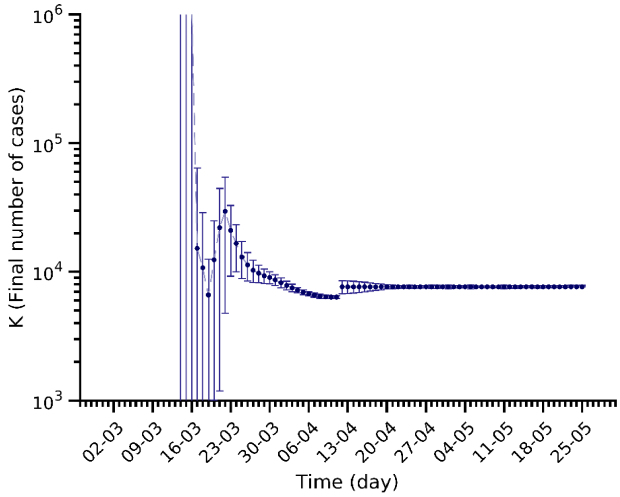


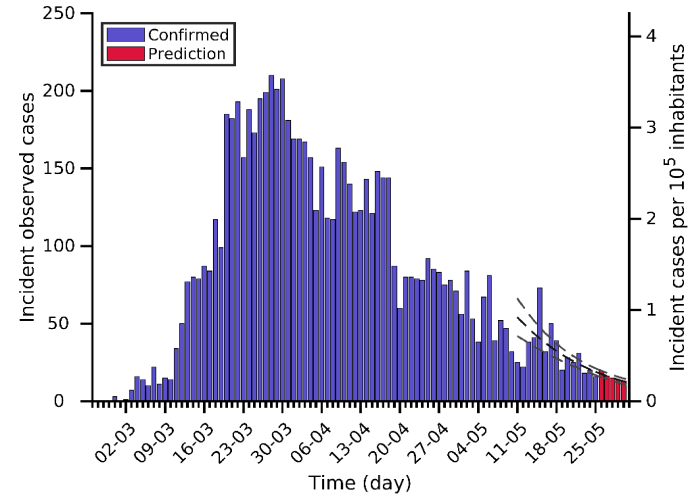
Lazio 25-05-2020. Population: 5.9M. Current cumulated incidence: 130/10⁵

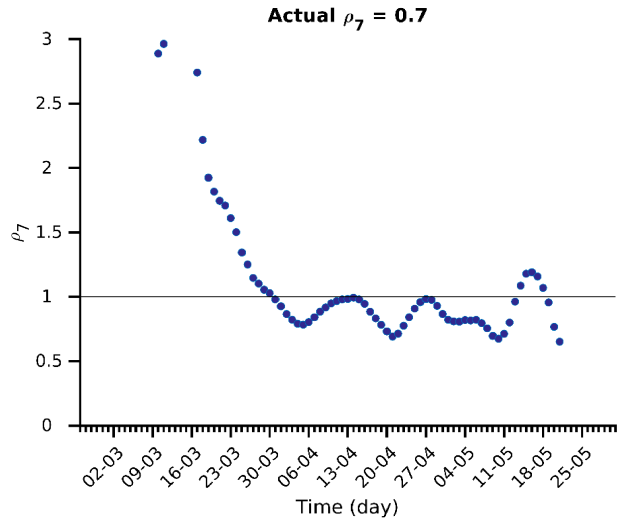


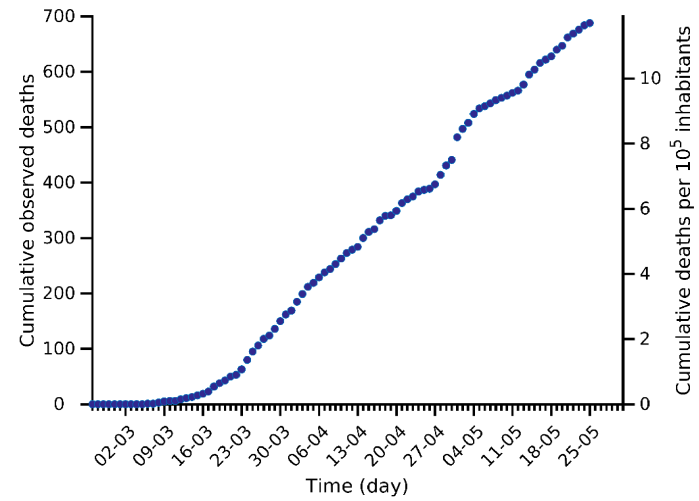


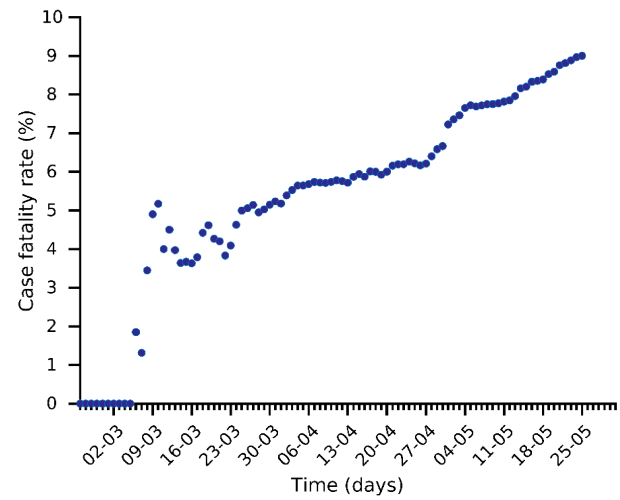




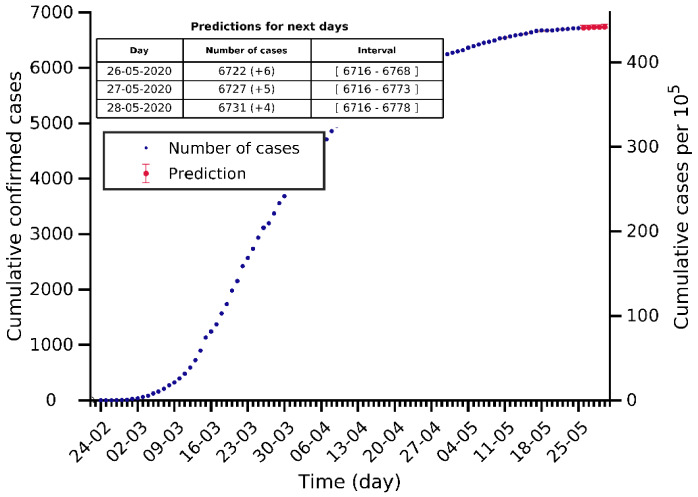


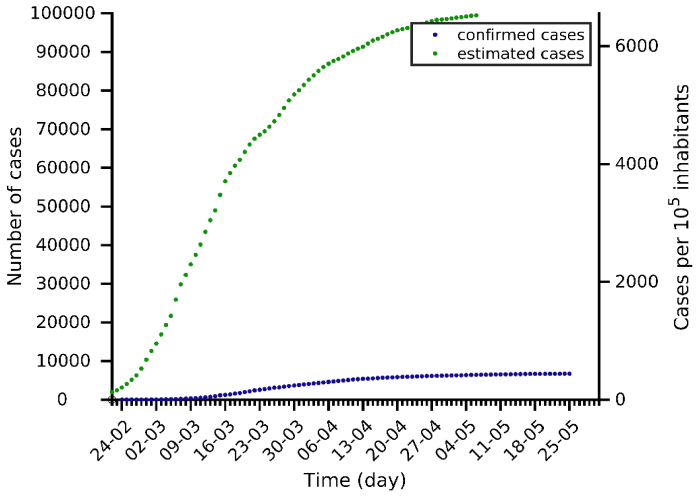


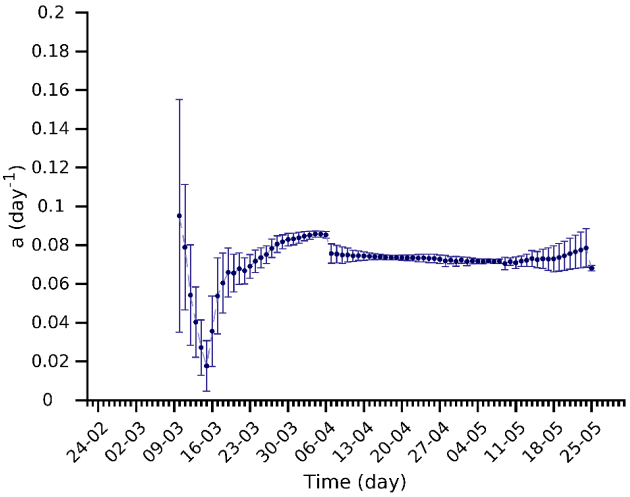


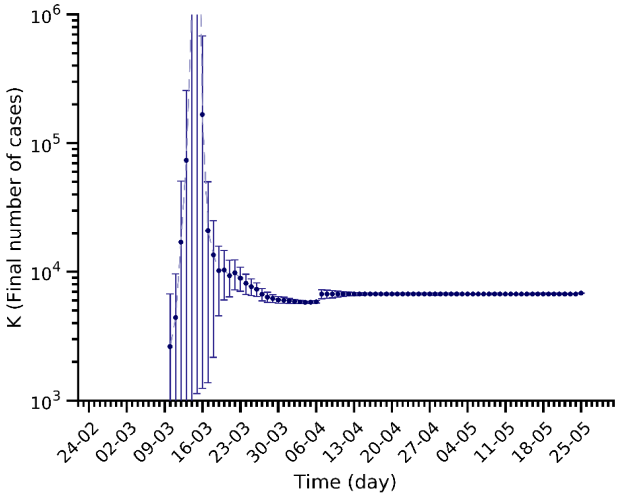


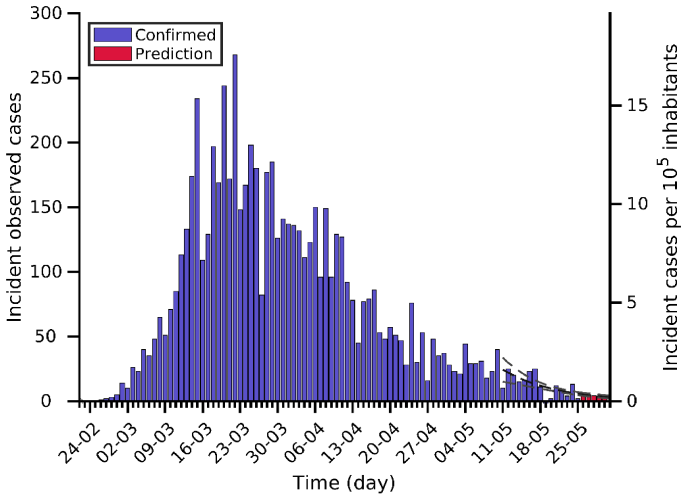
Marche 25-05-2020. Population: 1.5M. Current cumulated incidence: 440/10⁵

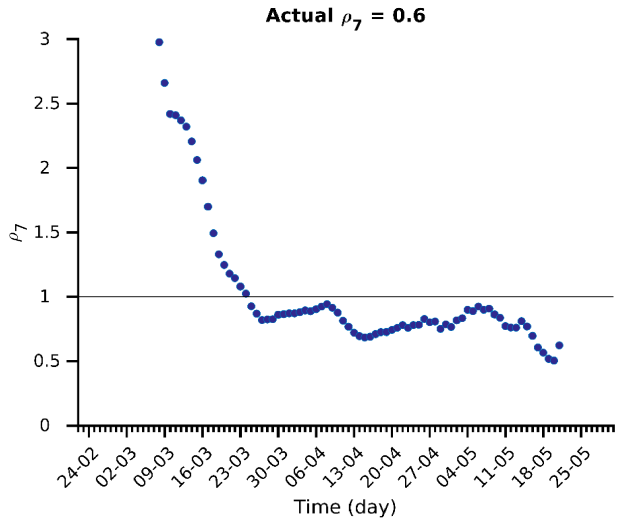


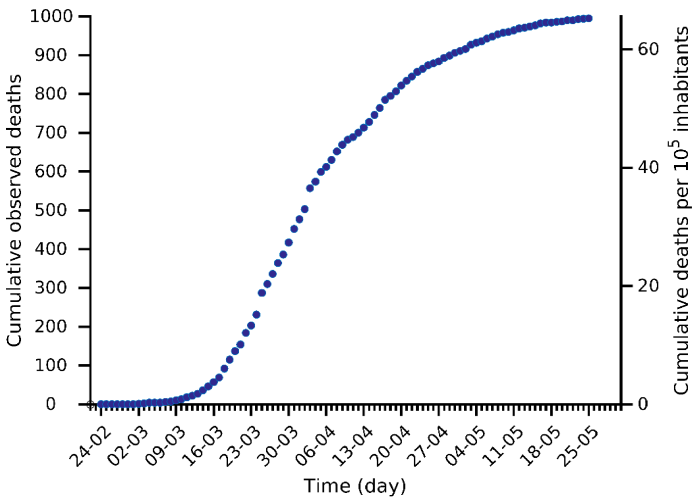


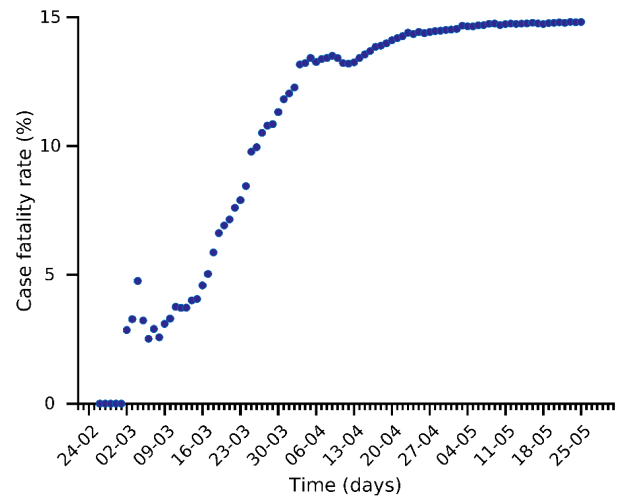




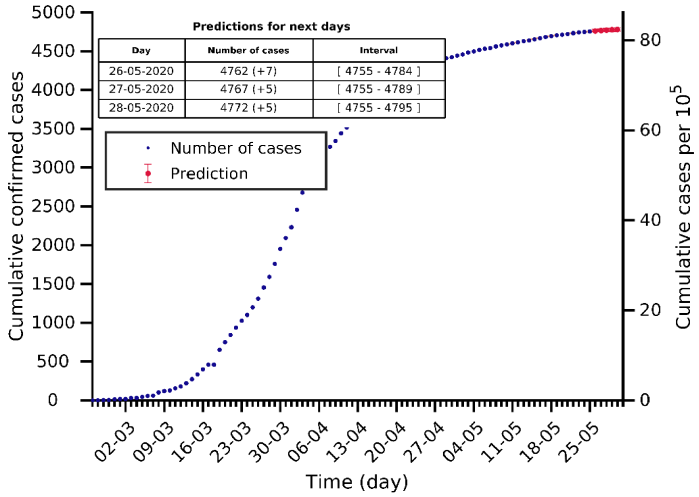


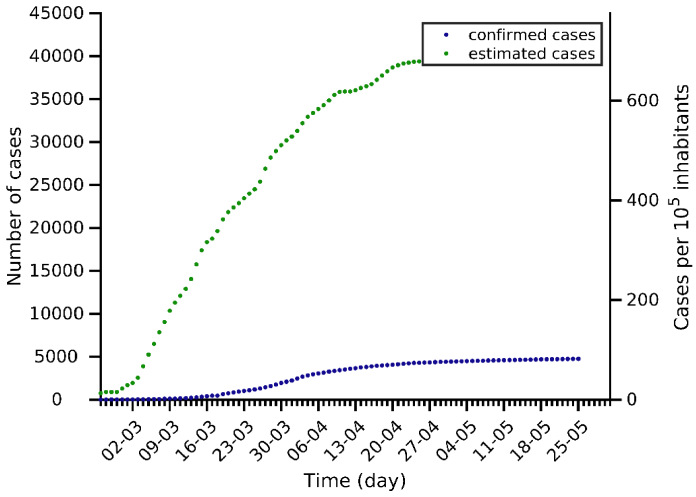


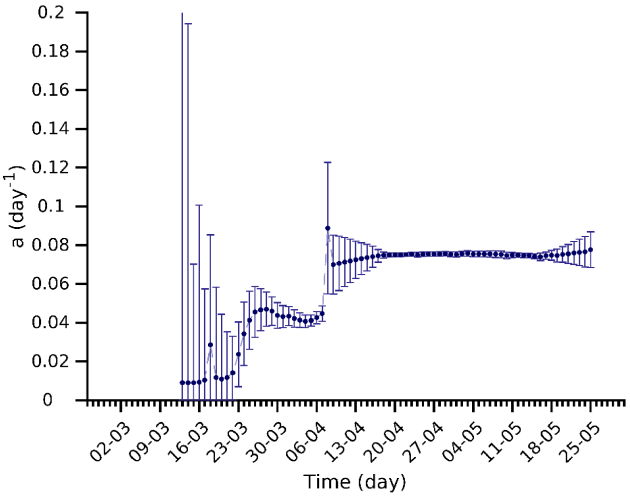


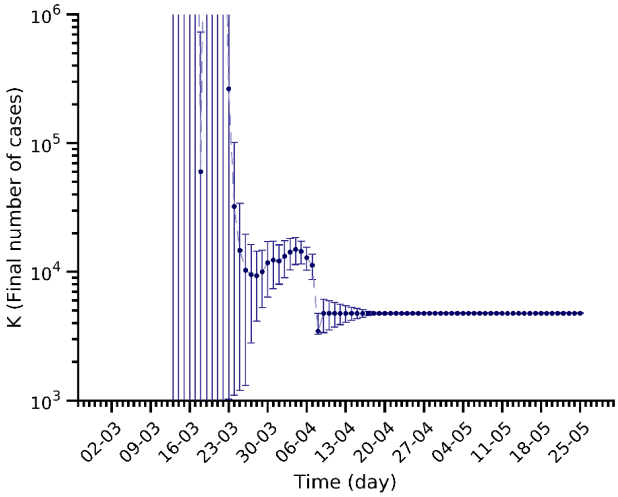


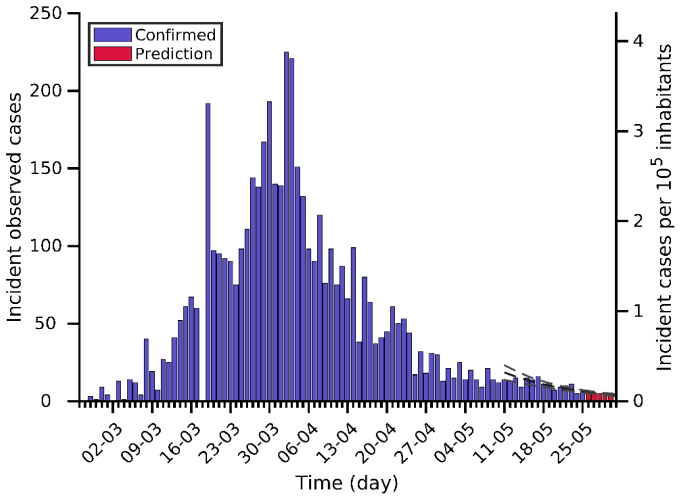
Campania 25-05-2020. Population: 5.8M. Current cumulated incidence: $82/10^5$

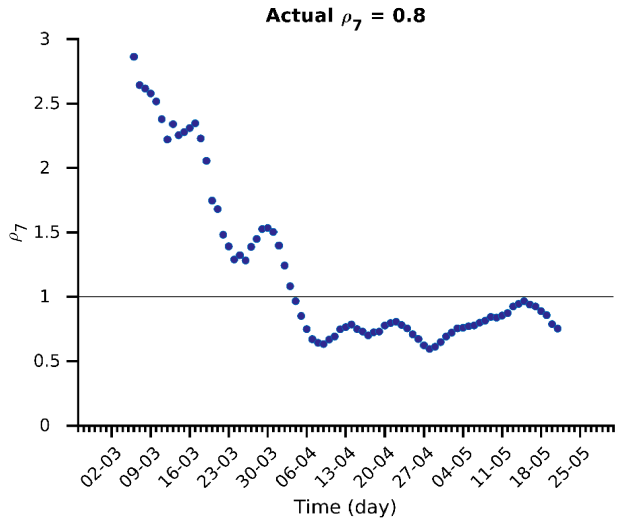


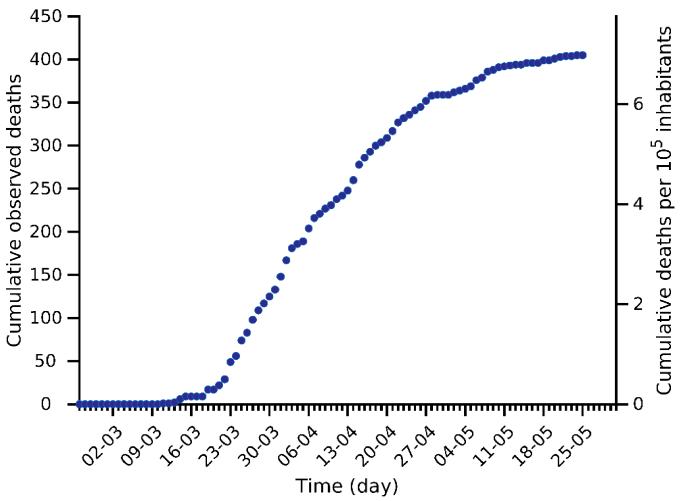


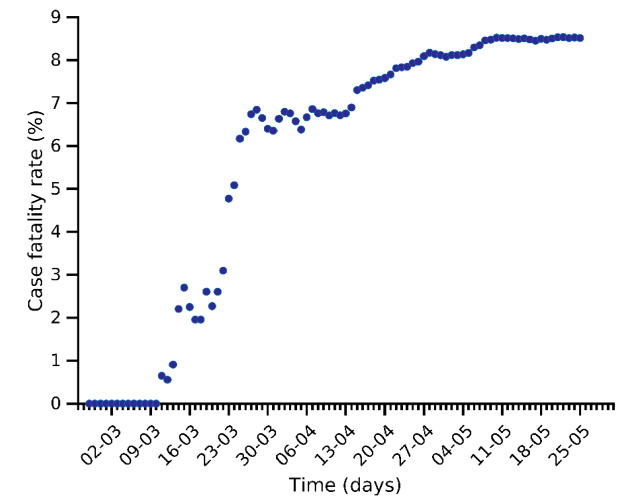




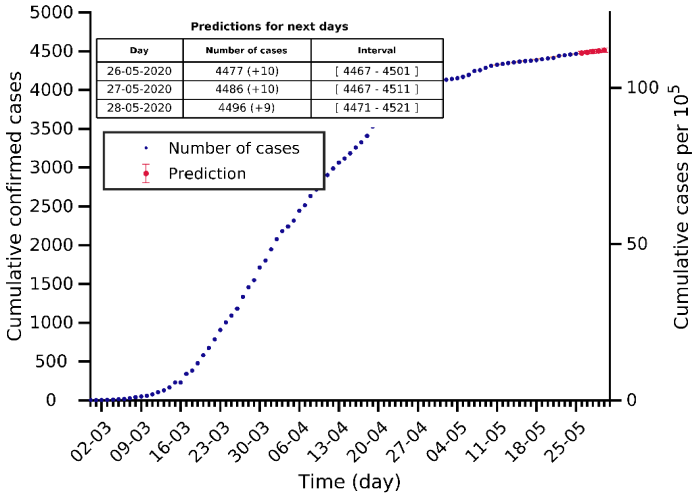


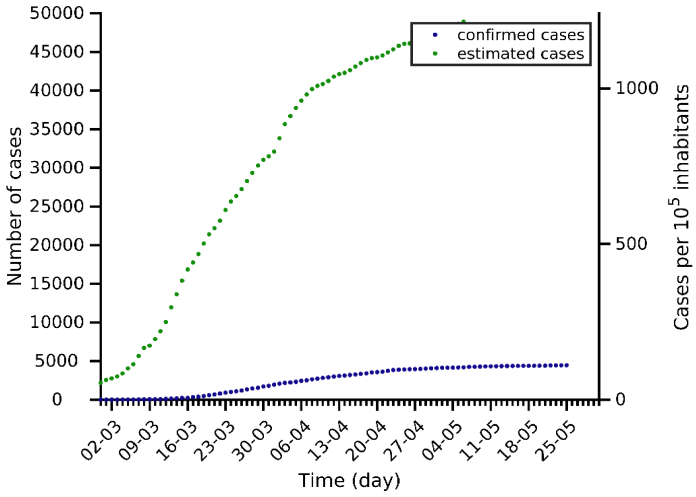


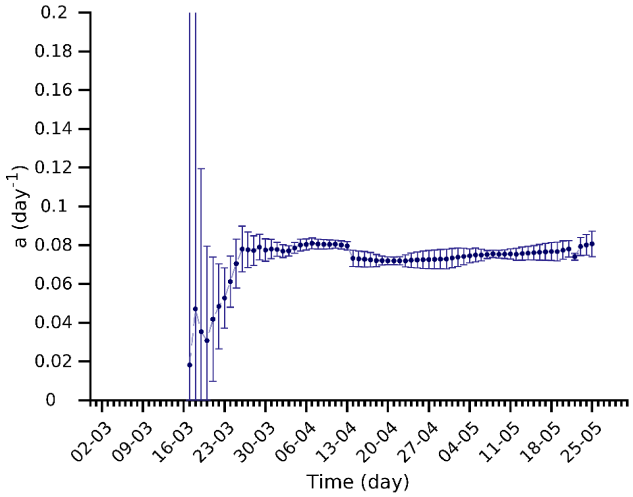


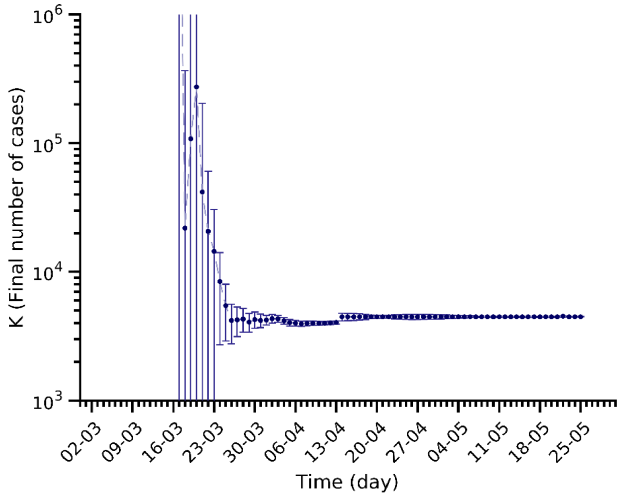


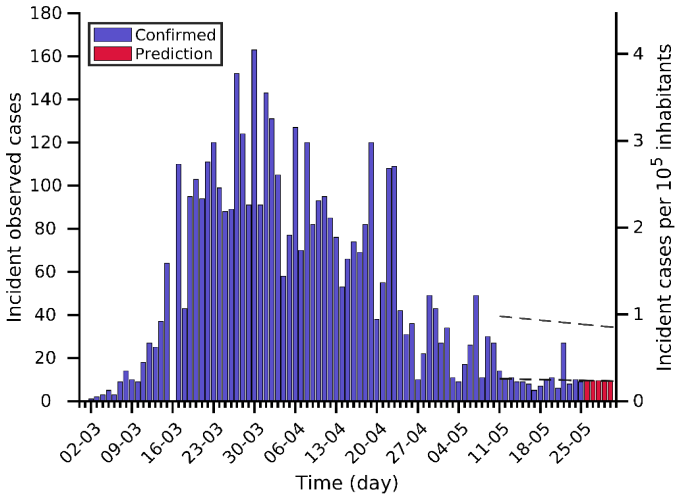
Puglia 25-05-2020. Population: 4.0M. Current cumulated incidence: 111/10⁵

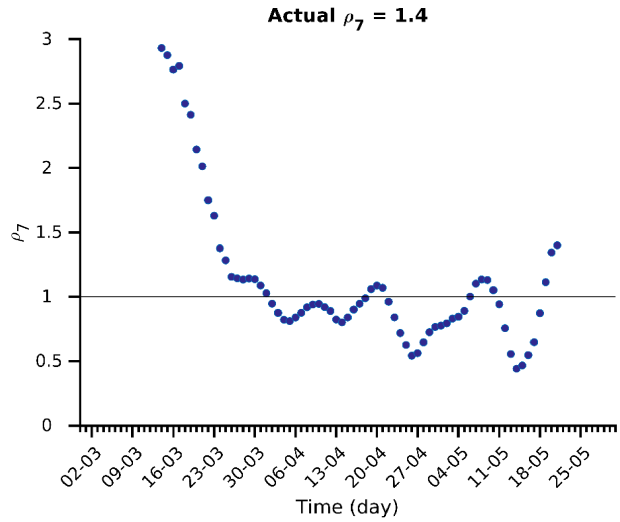


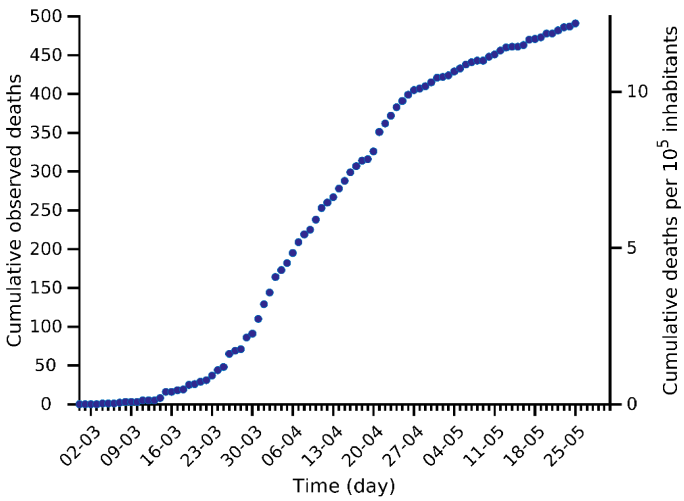


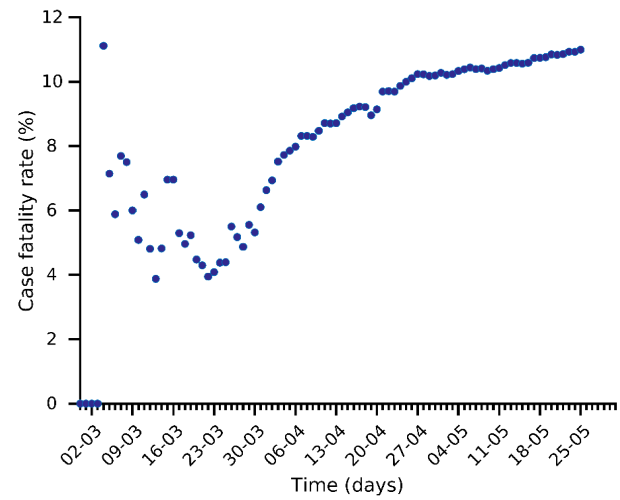




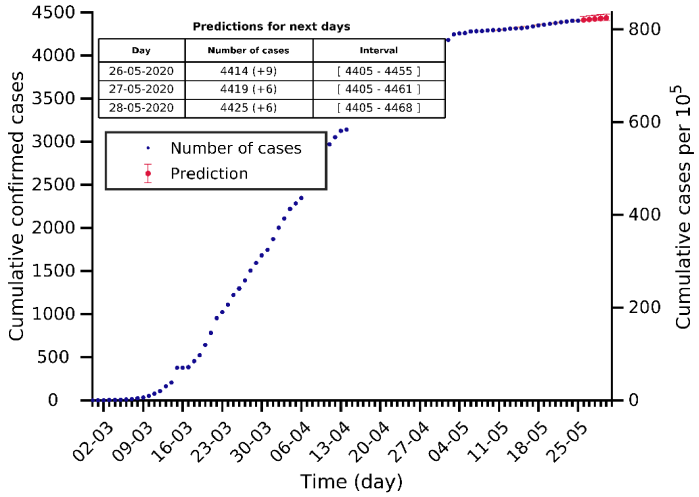


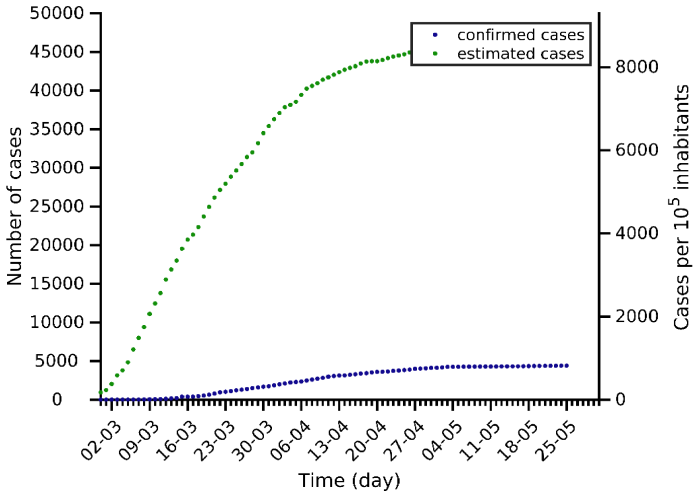


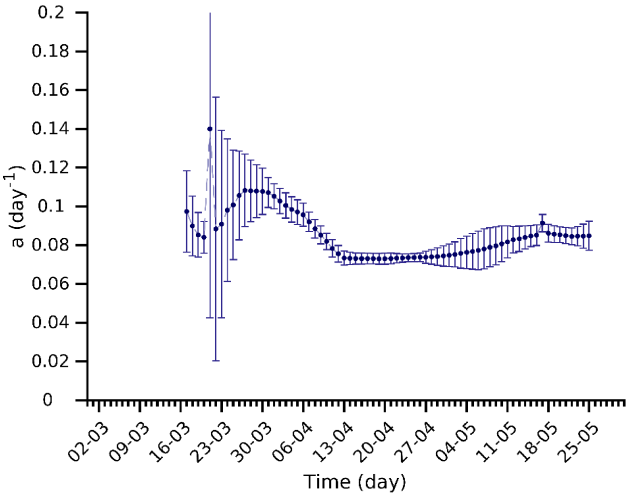


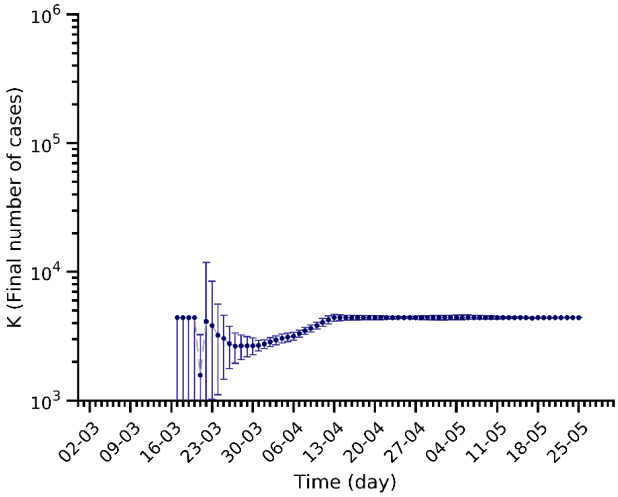


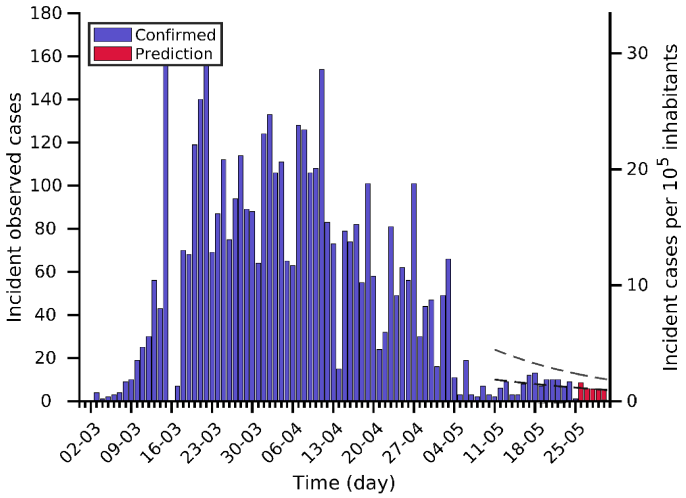
Trento 25-05-2020. Population: 0.5M. Current cumulated incidence: 819/10^5

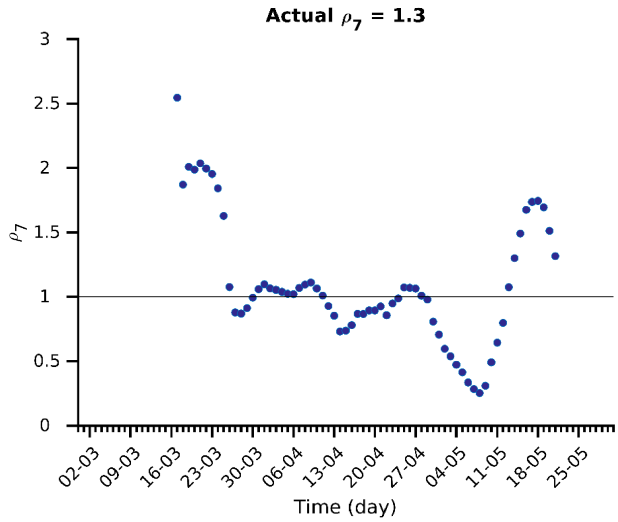


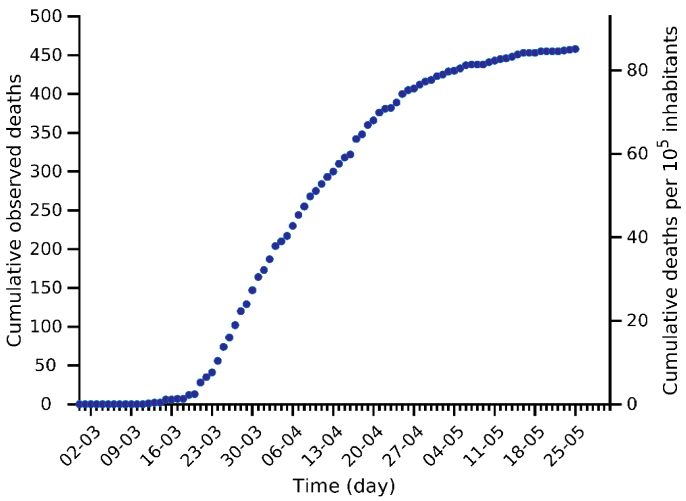


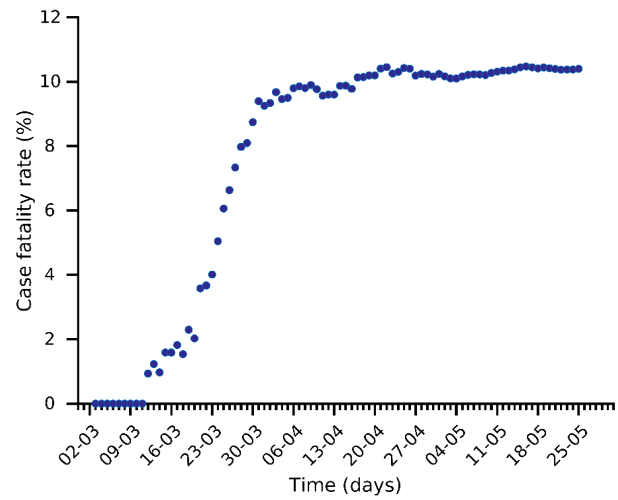




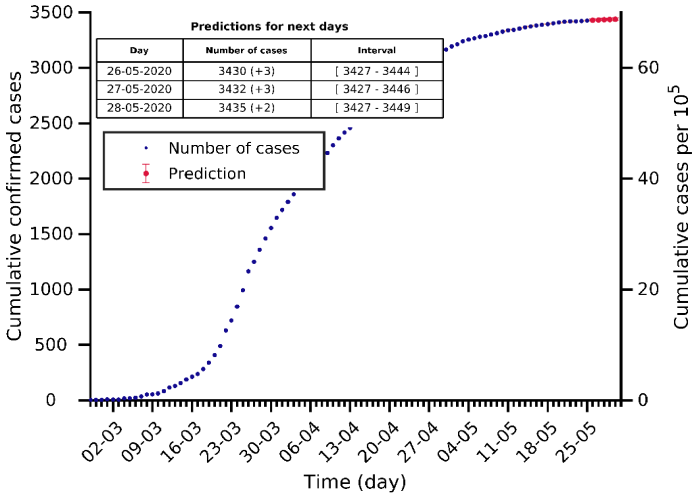


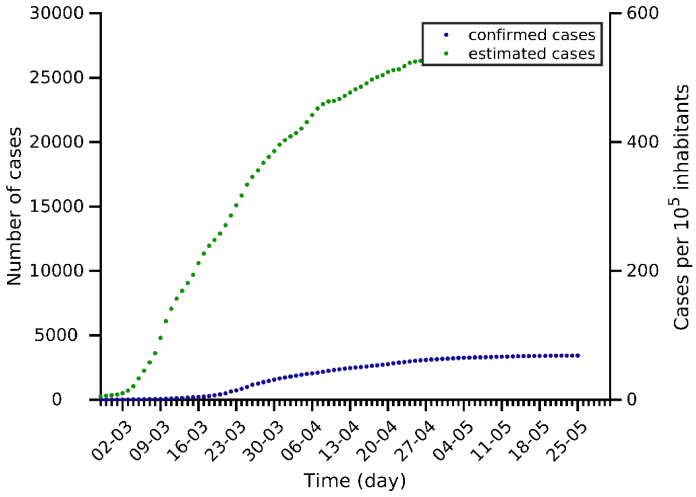


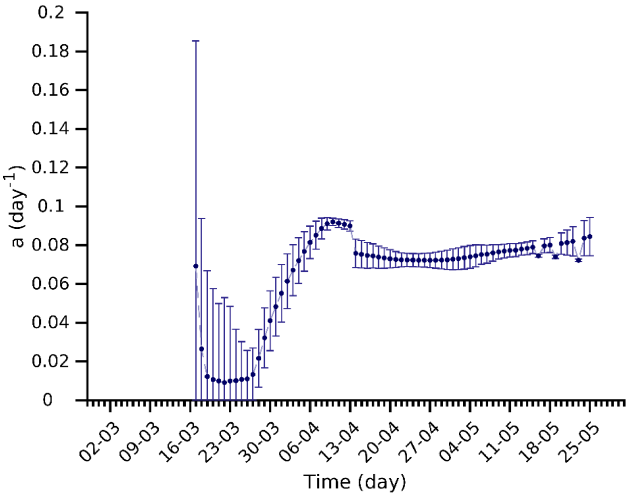


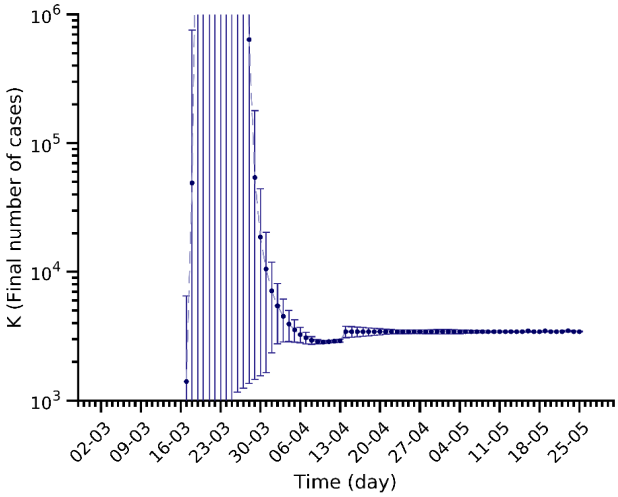


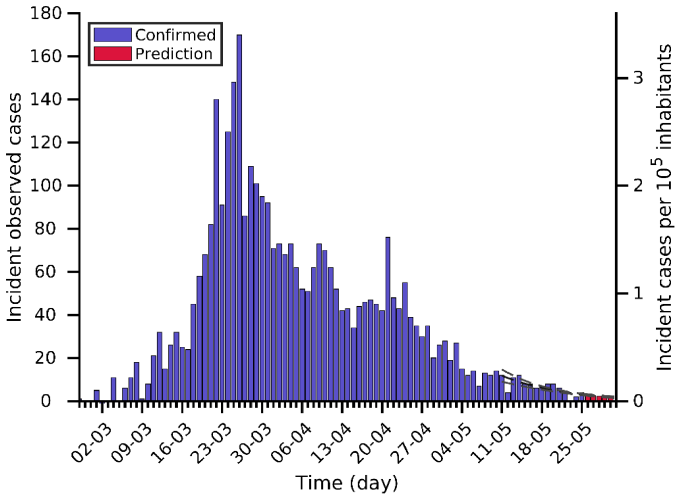
Sicilia 25-05-2020. Population: 5.0M. Current cumulated incidence: 69/10^5

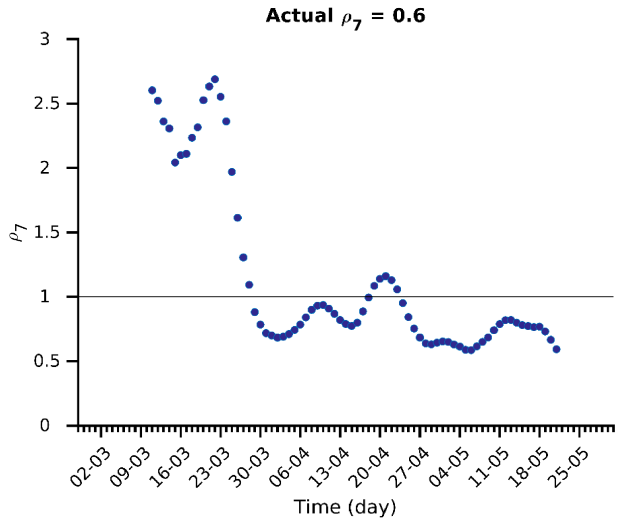


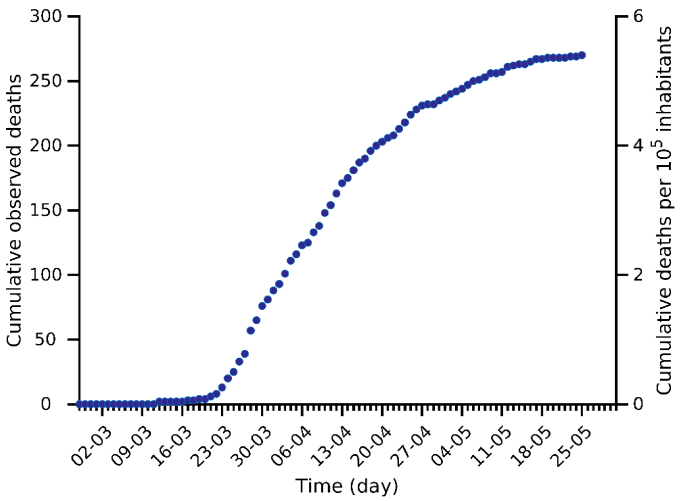


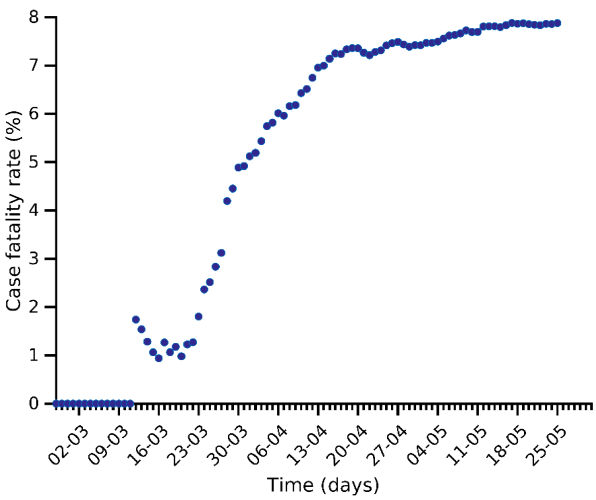




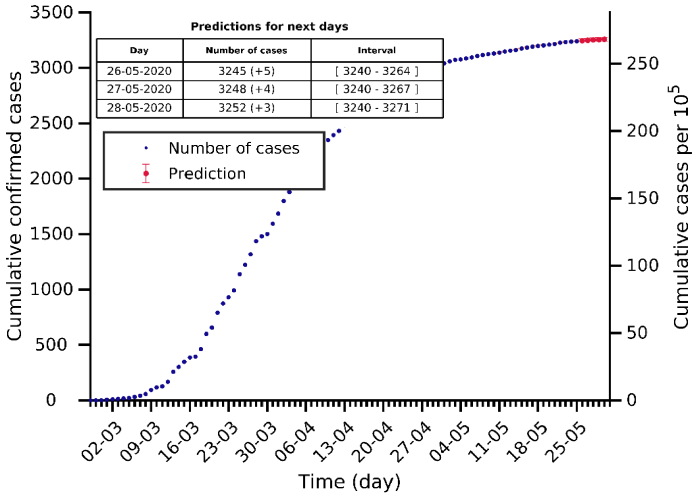


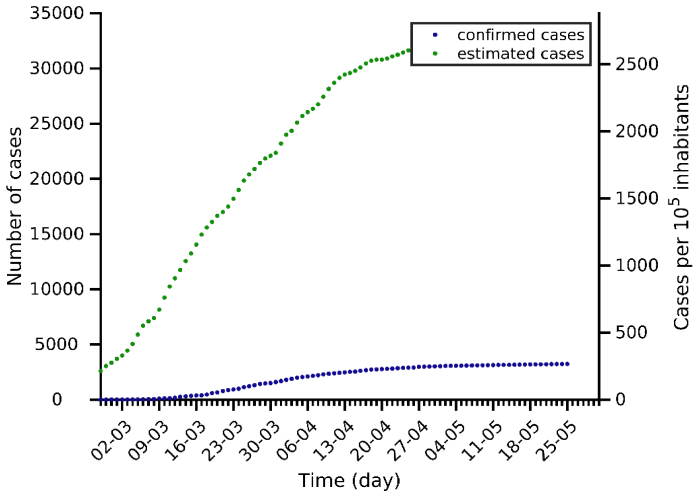


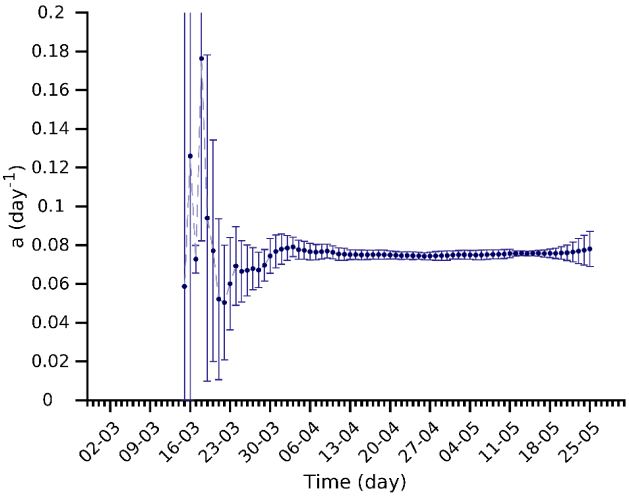


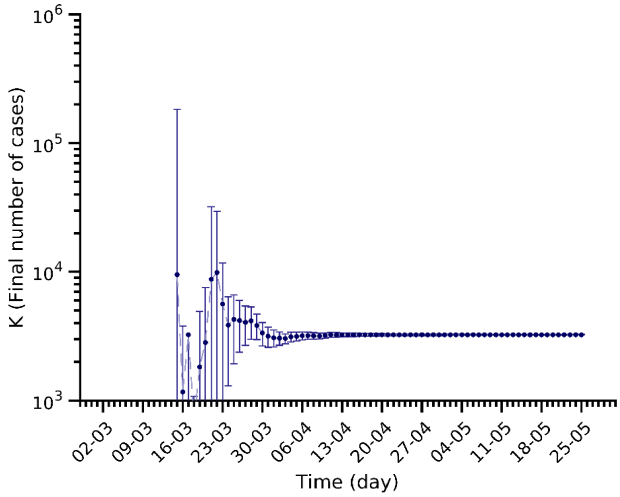


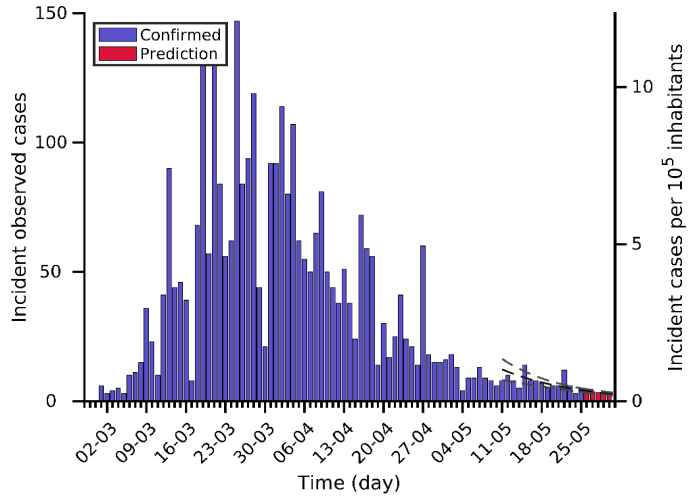
Friuli Venezia Giulia 25-05-2020. Population: 1.2M. Current cumulated incidence: 2

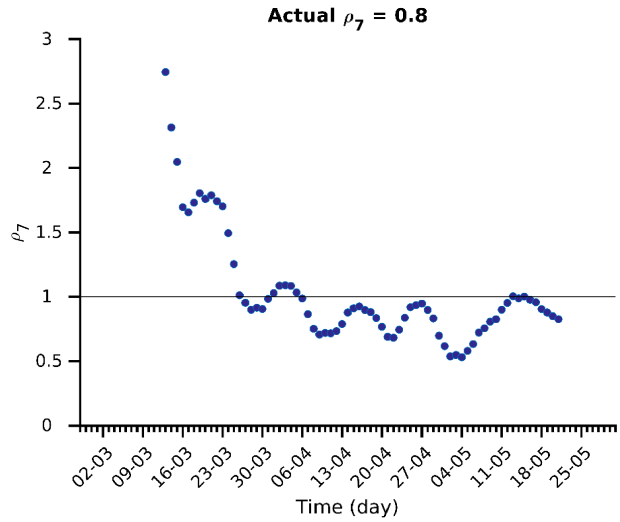


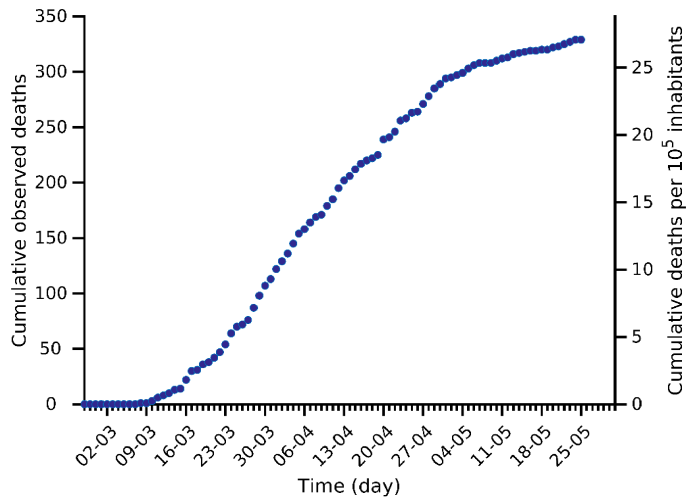


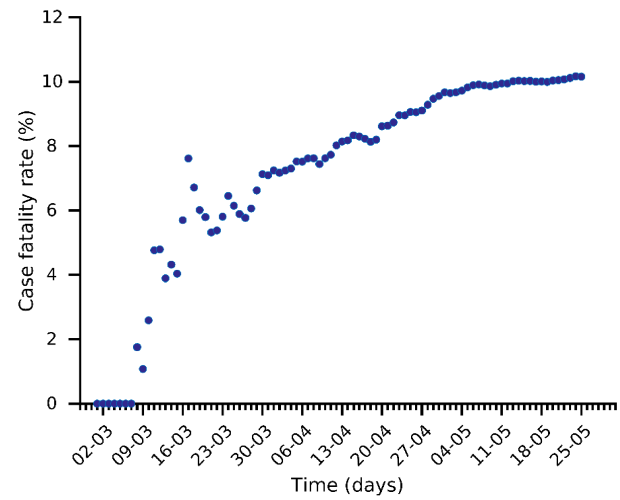




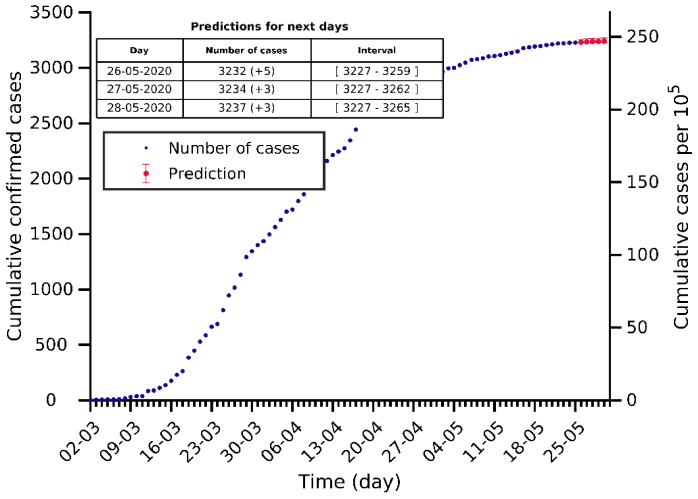


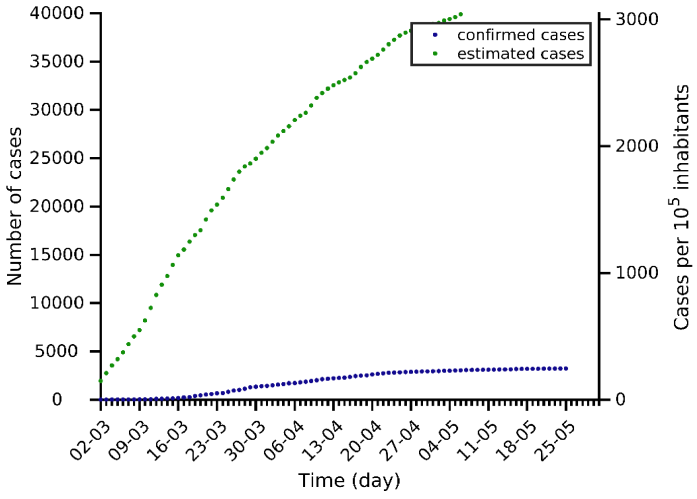


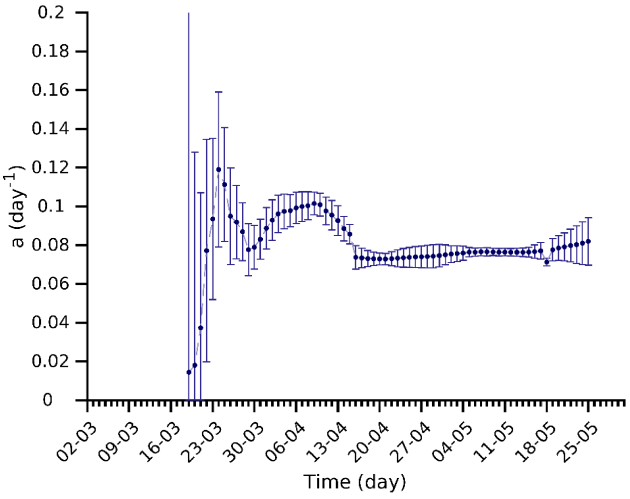


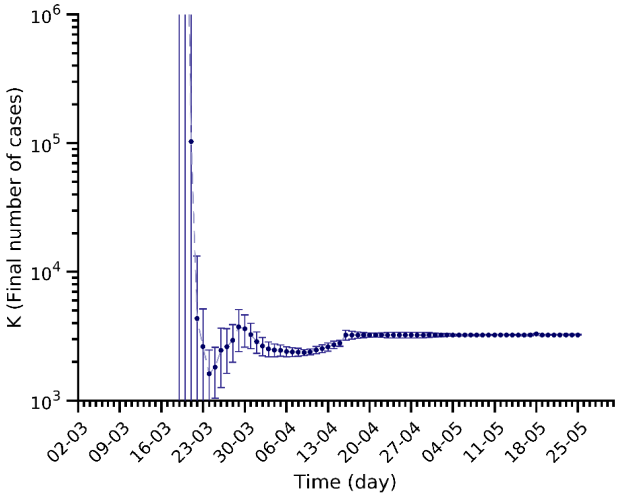


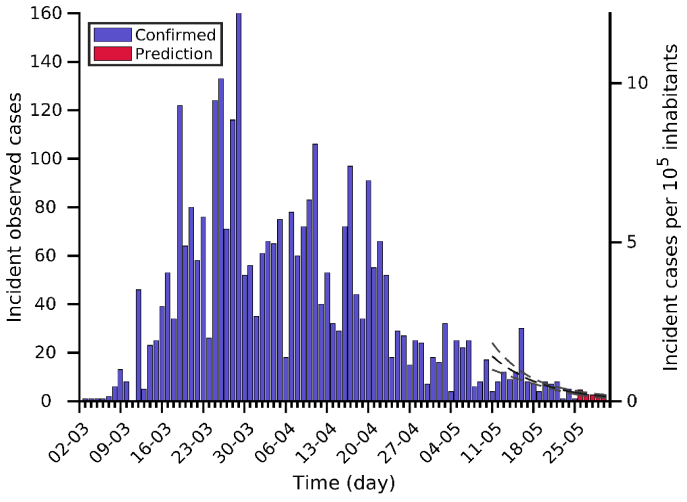
Abruzzo 25-05-2020. Population: 1.3M. Current cumulated incidence: 246/10^5

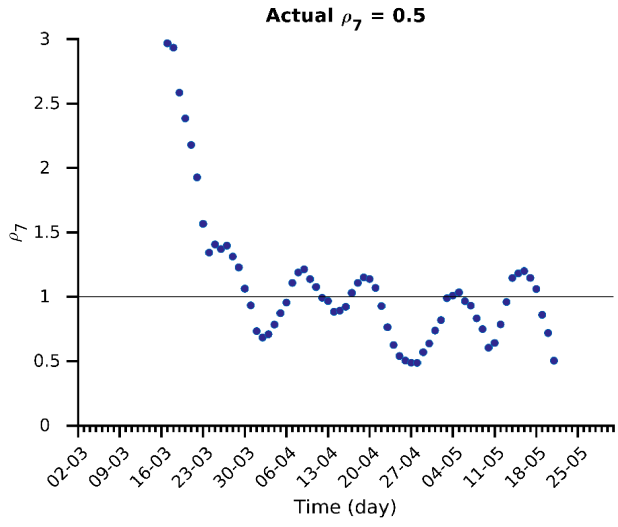


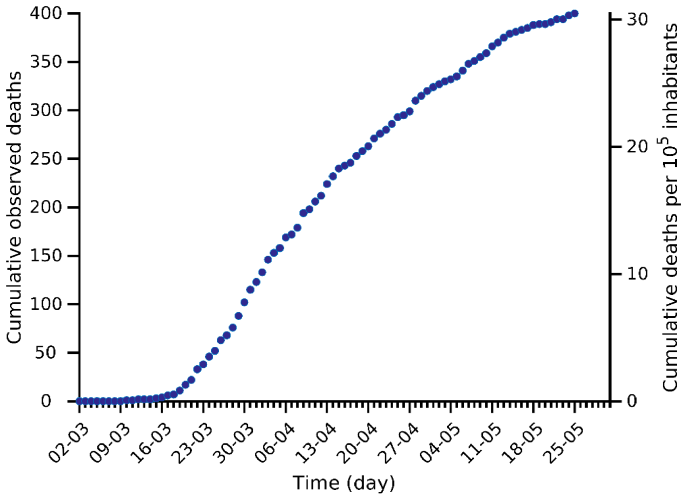


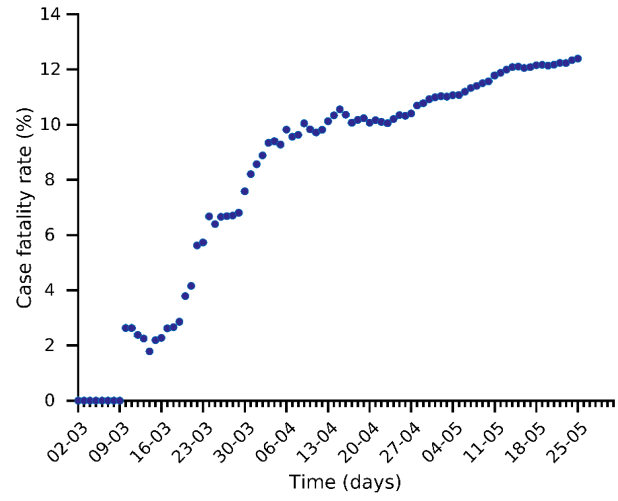




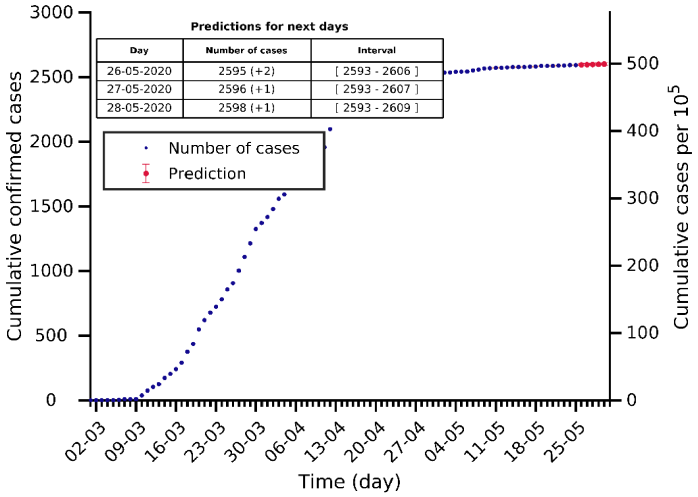


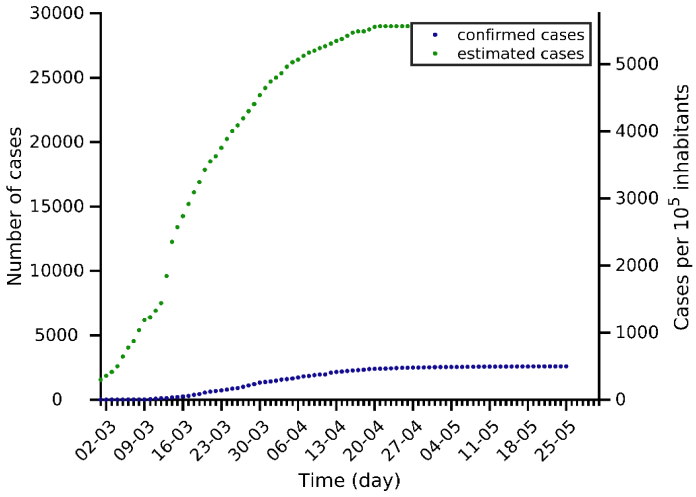


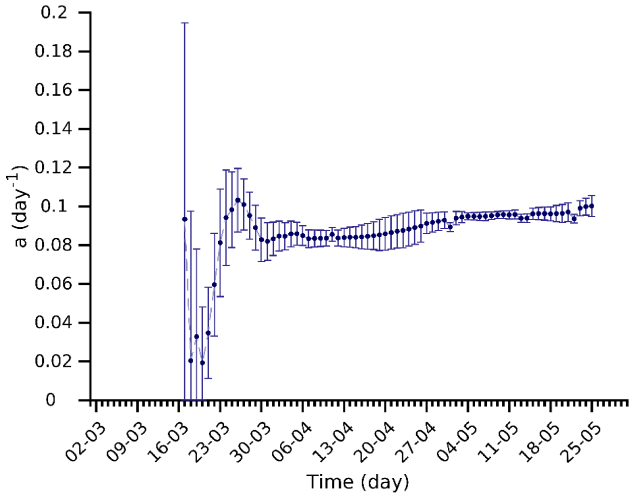


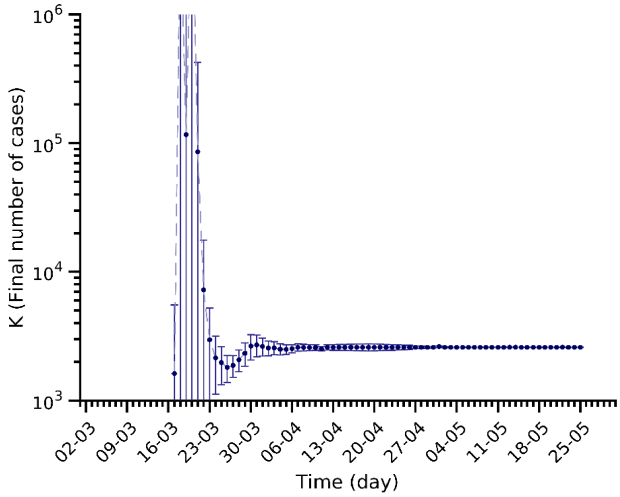


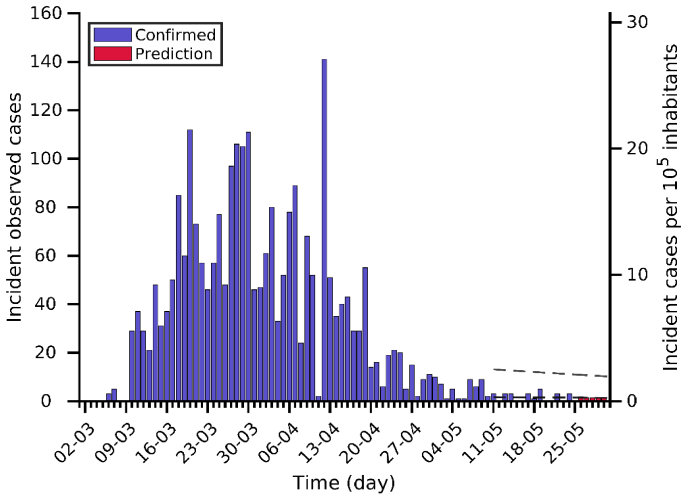
Bolzano 25-05-2020. Population: 0.5M. Current cumulated incidence: 498/10^5

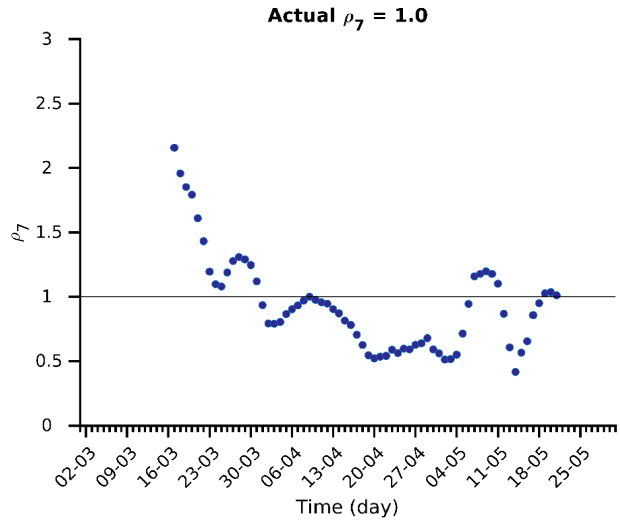


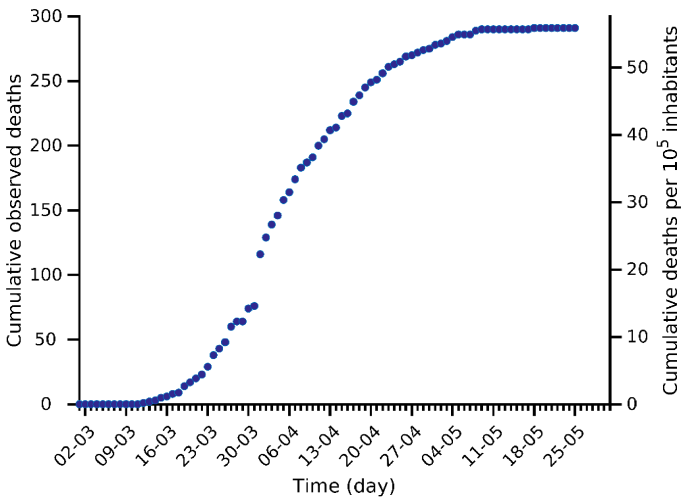


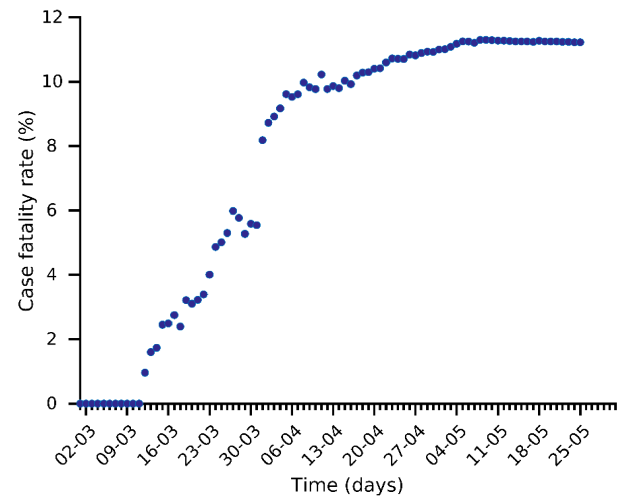




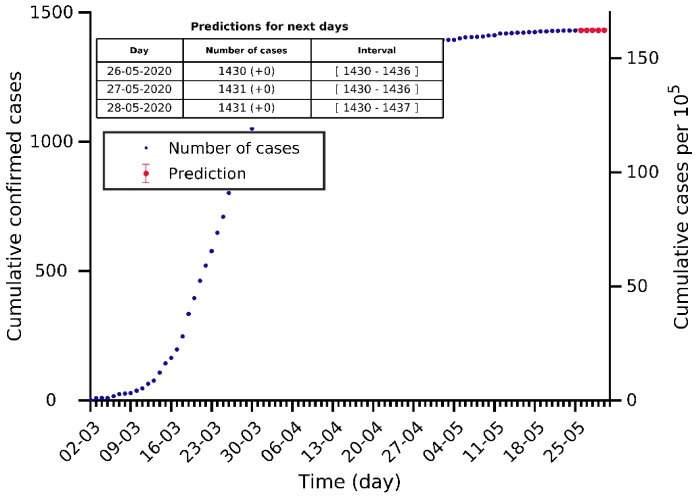


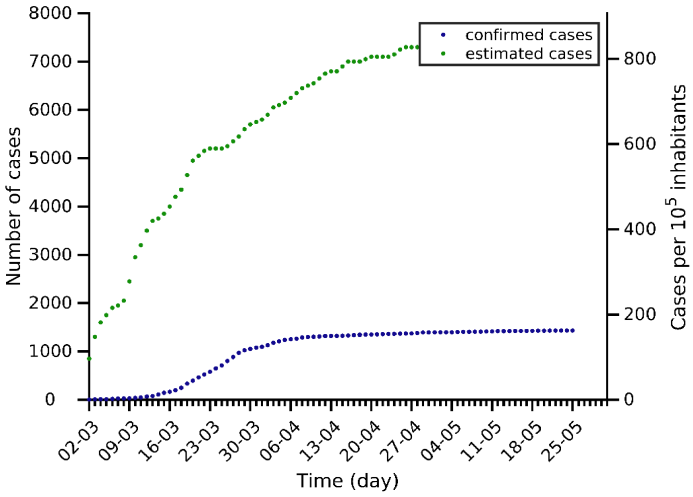


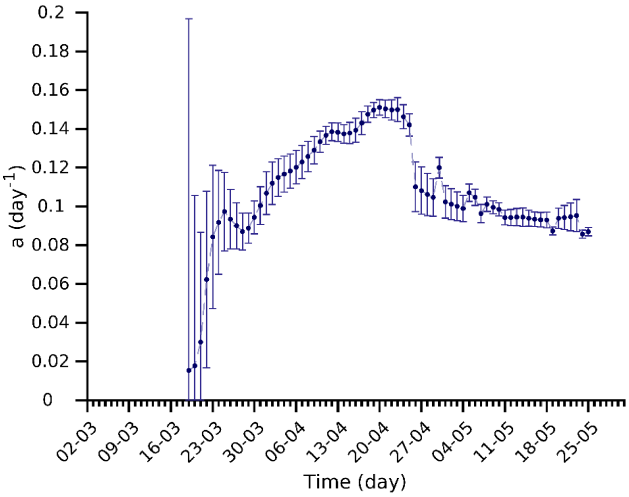


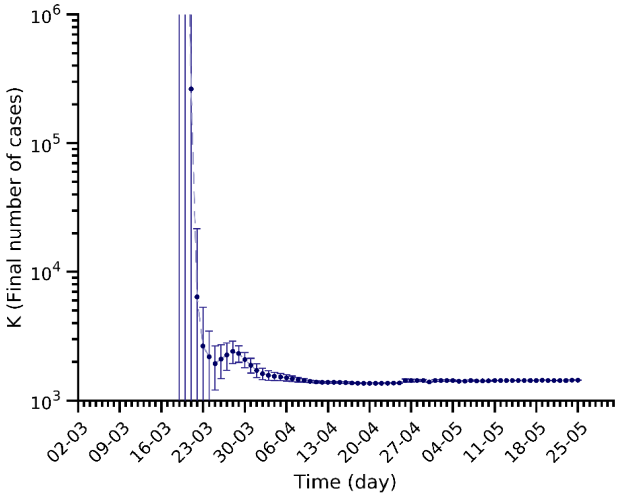


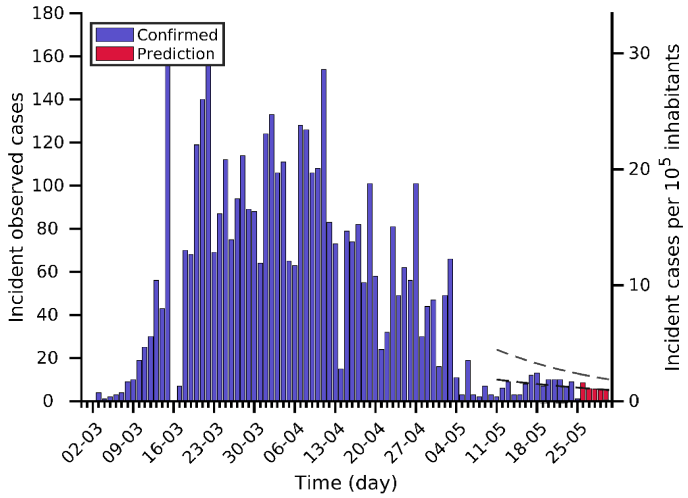
Umbria 25-05-2020. Population: 0.9M. Current cumulated incidence: $162/10^5$

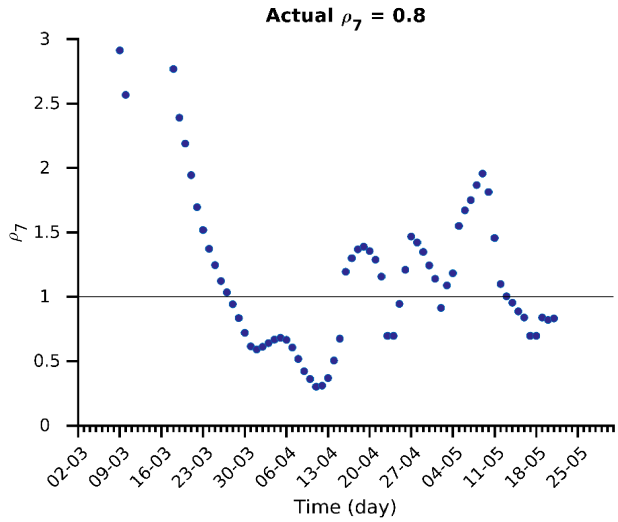


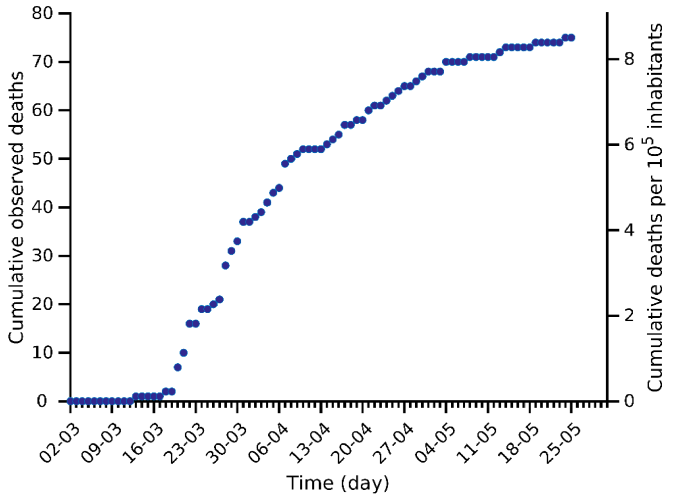


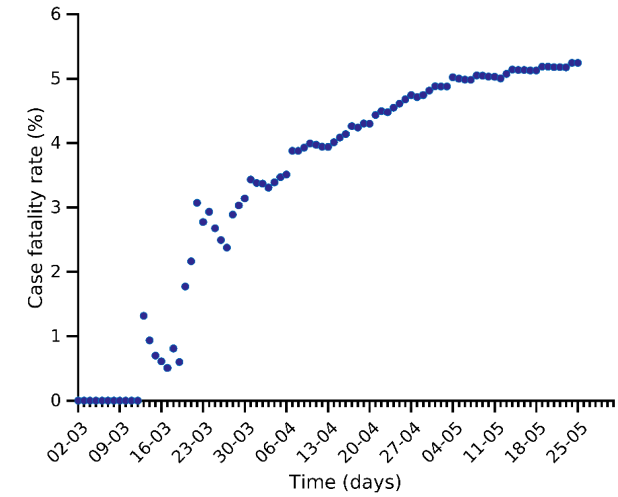




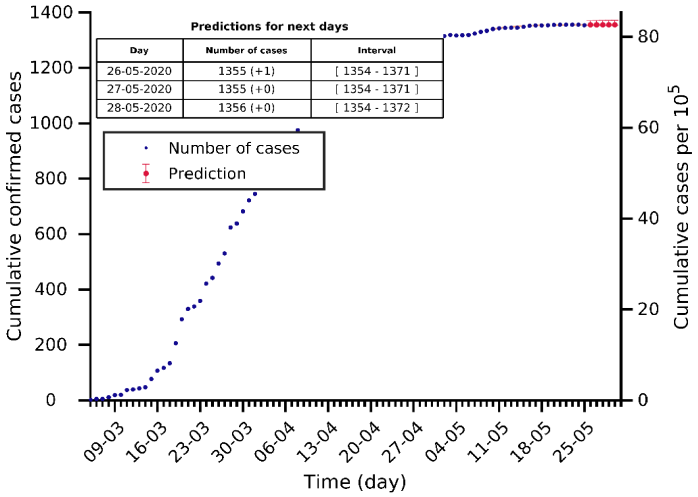


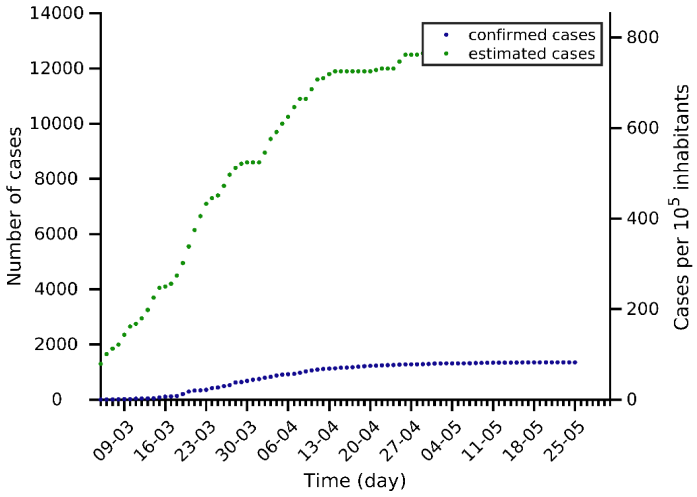


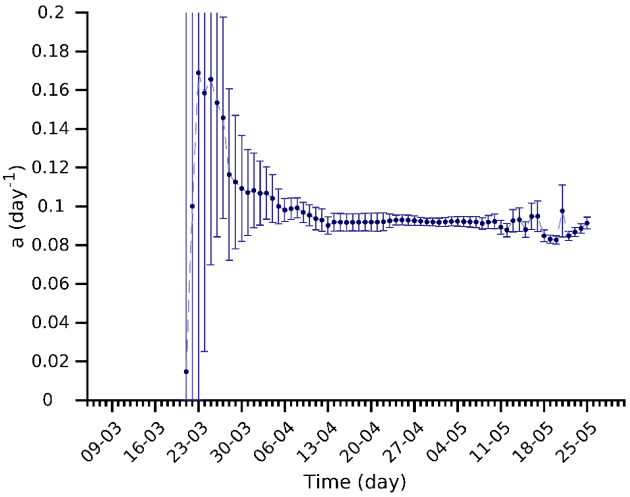


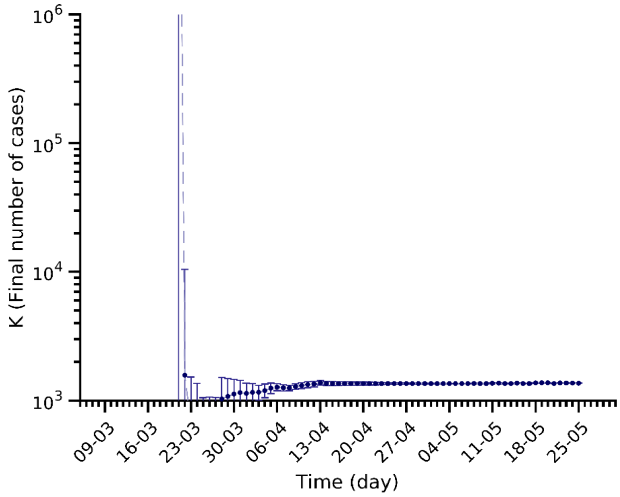


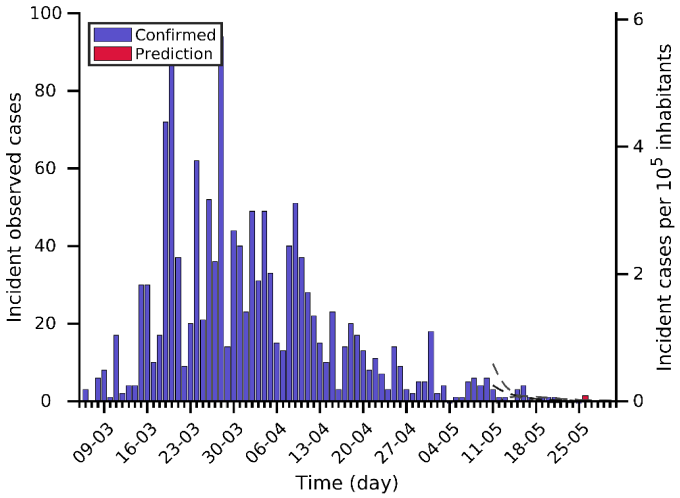
Sardegna 25-05-2020. Population: 1.6M. Current cumulated incidence: 83/10^5

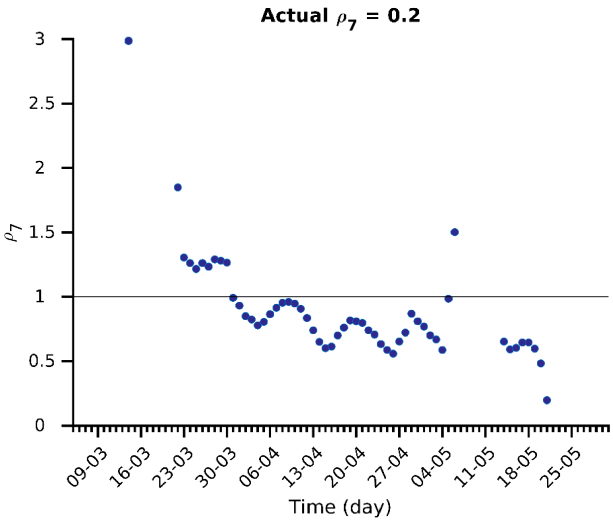


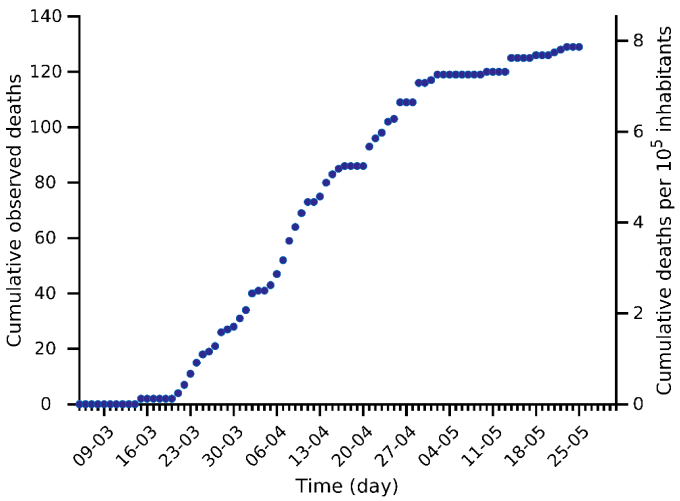


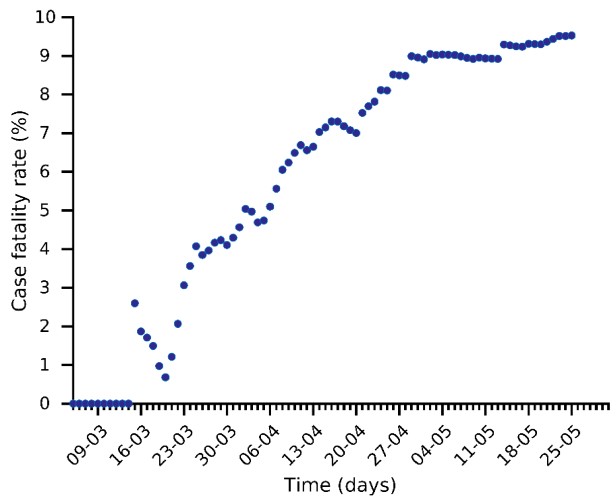




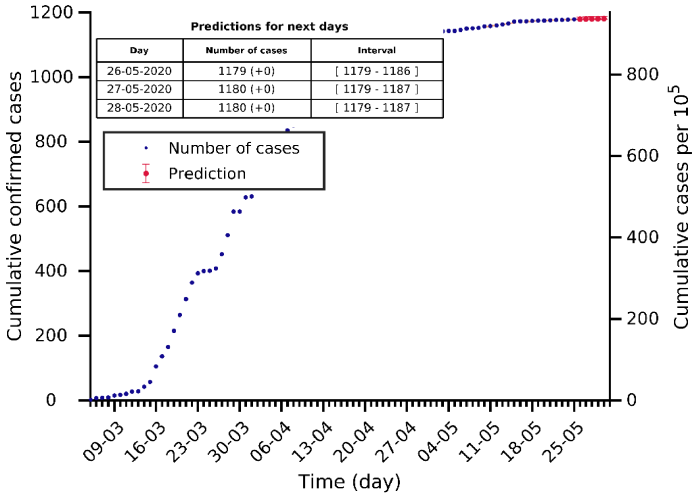


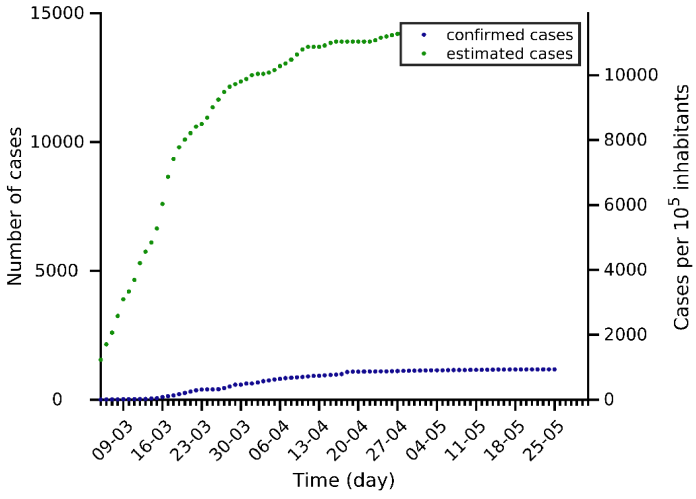


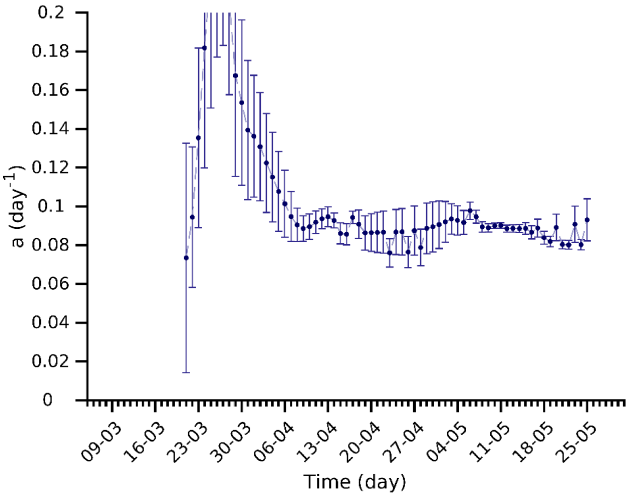


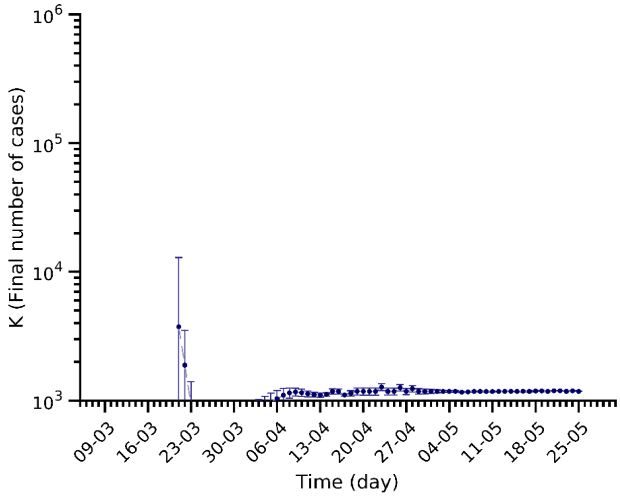


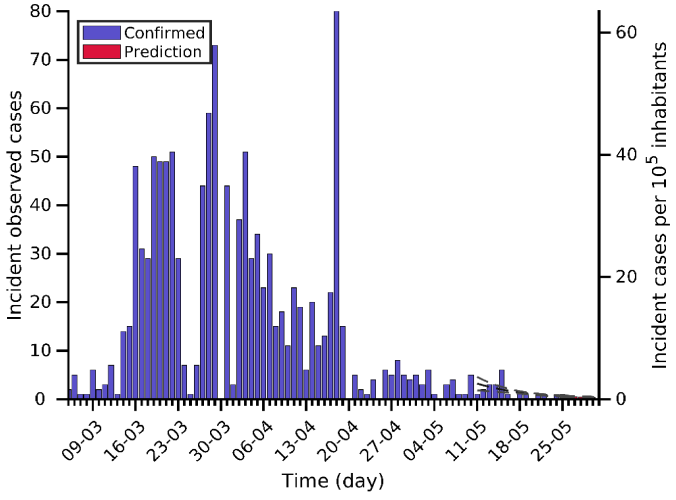
Valle dAosta 25-05-2020. Population: 0.1M. Current cumulated incidence: 936/10^5

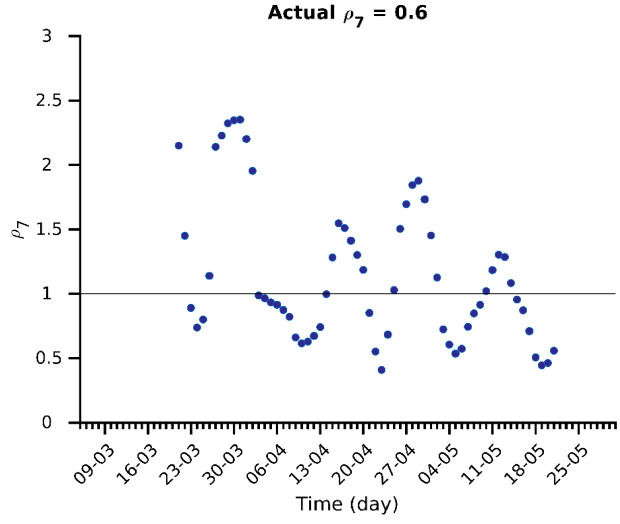


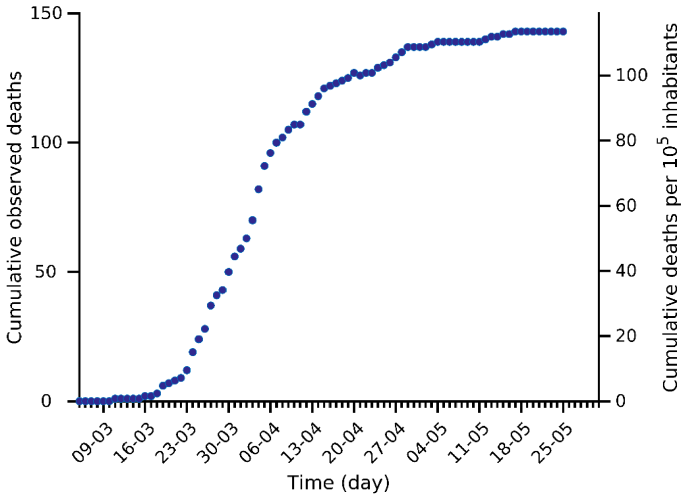


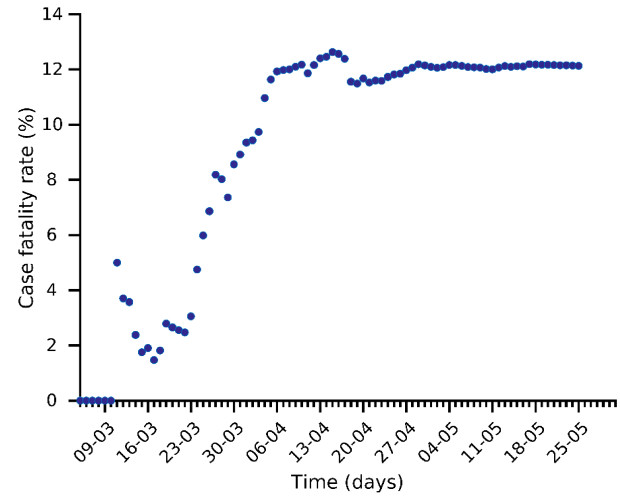




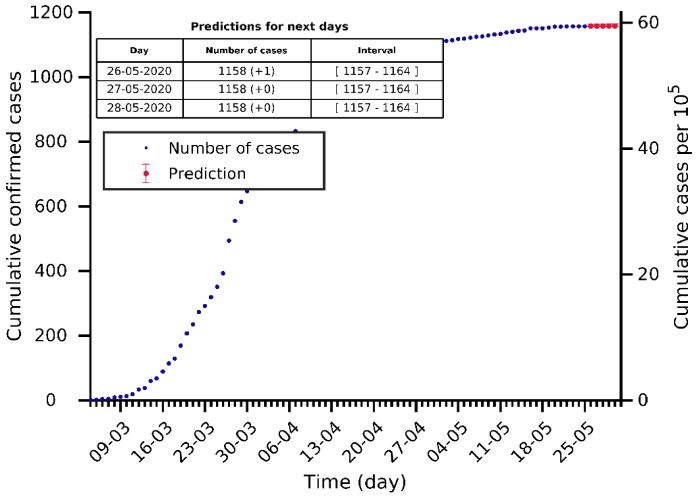


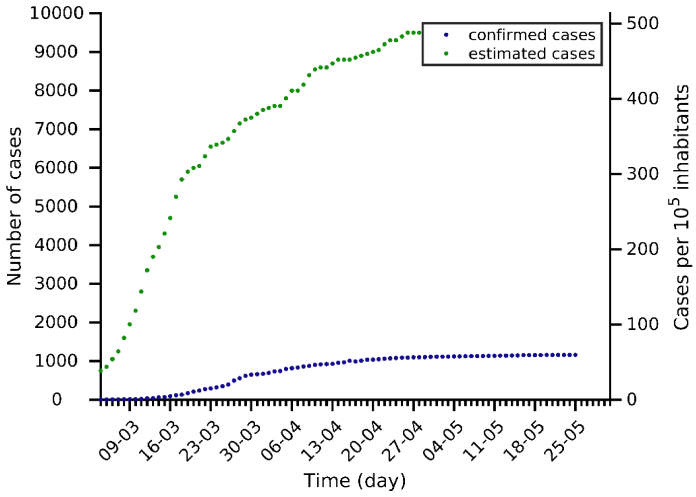


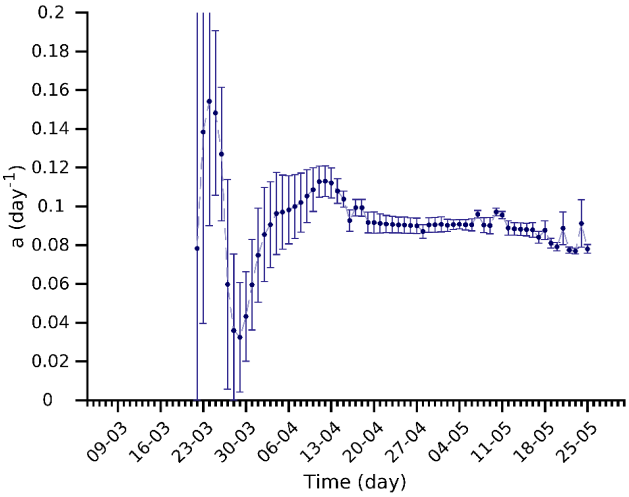


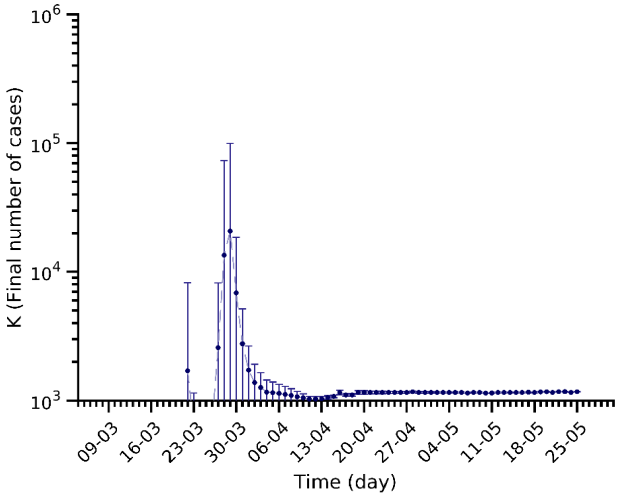


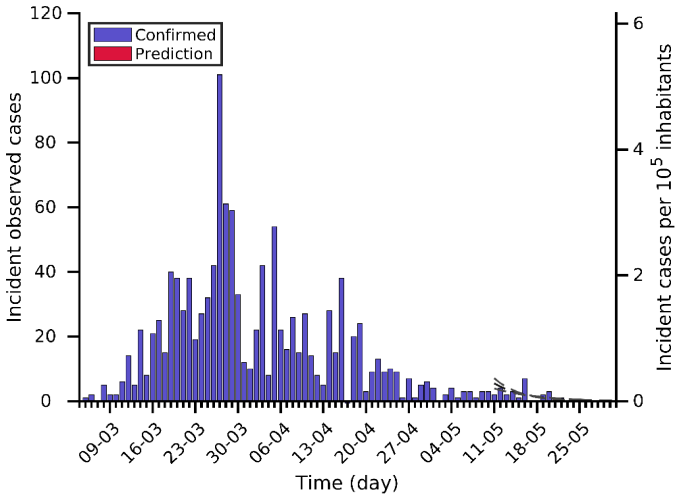
Calabria 25-05-2020. Population: 1.9M. Current cumulated incidence: 59/10^5

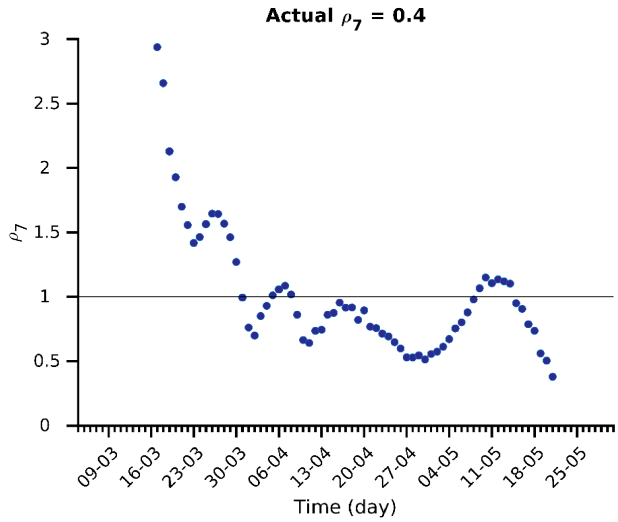


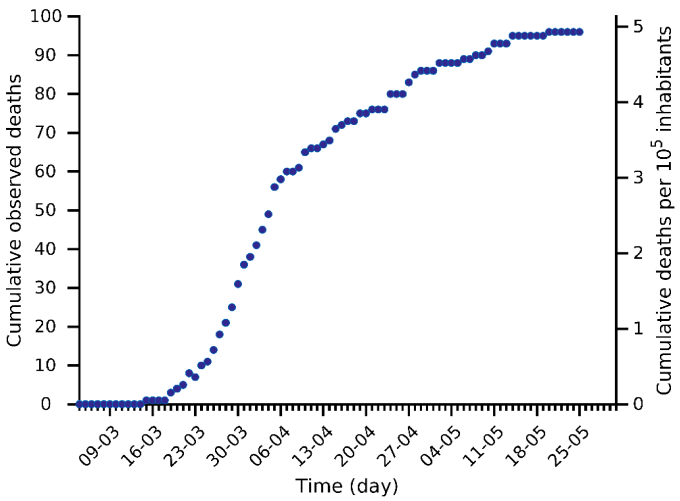


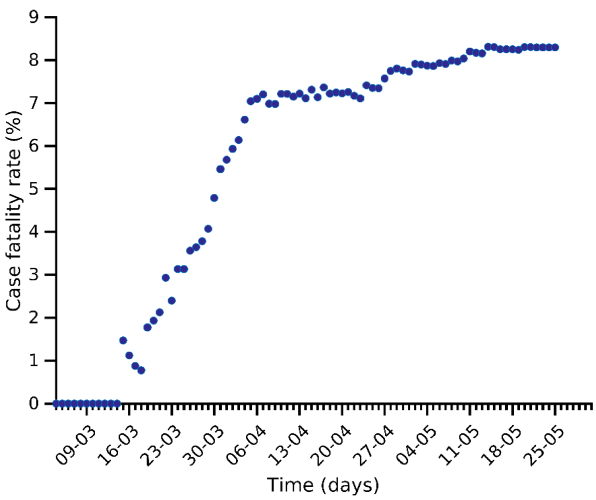




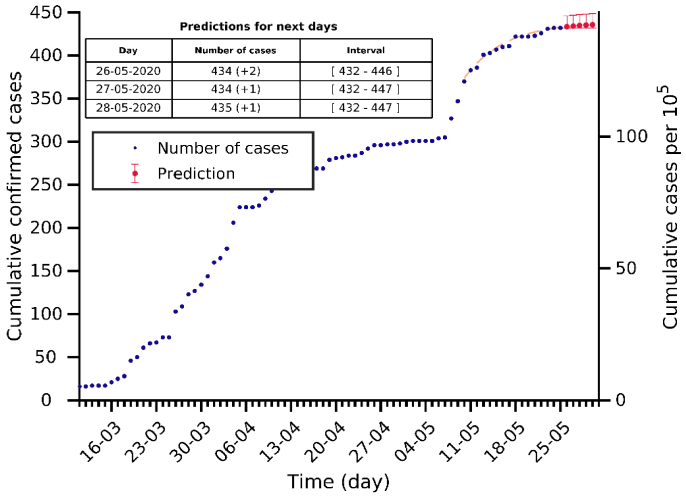


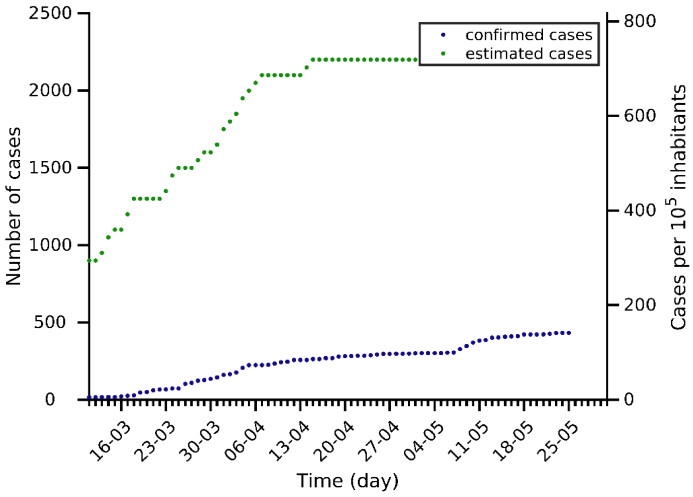


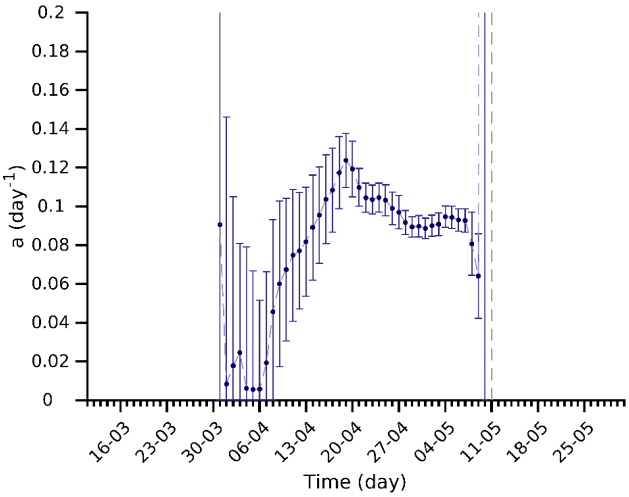


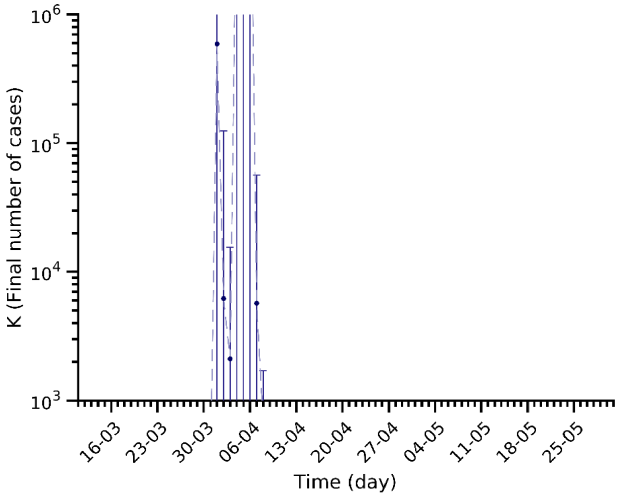


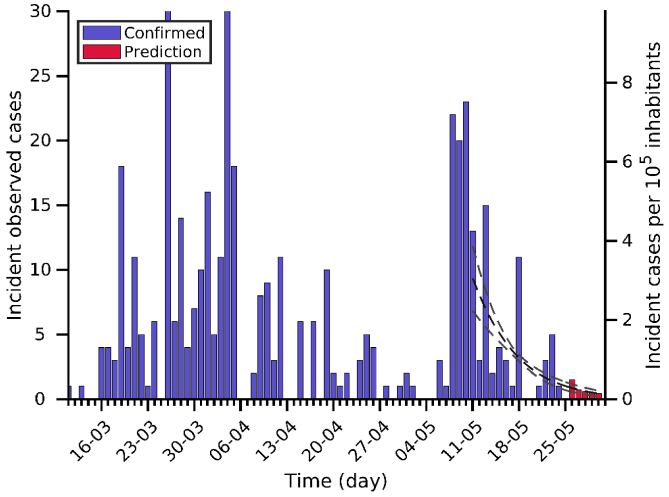
Molise 25-05-2020. Population: 0.3M. Current cumulated incidence: 141/10⁵

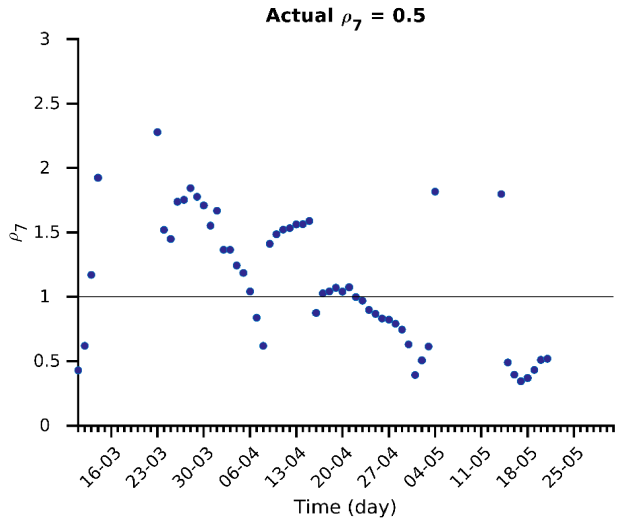


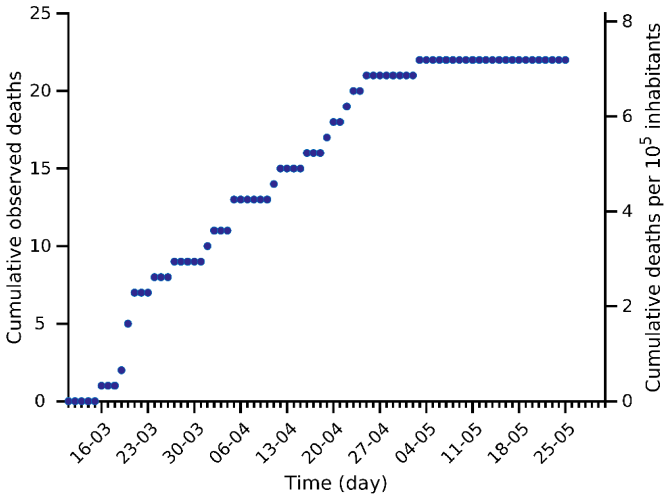


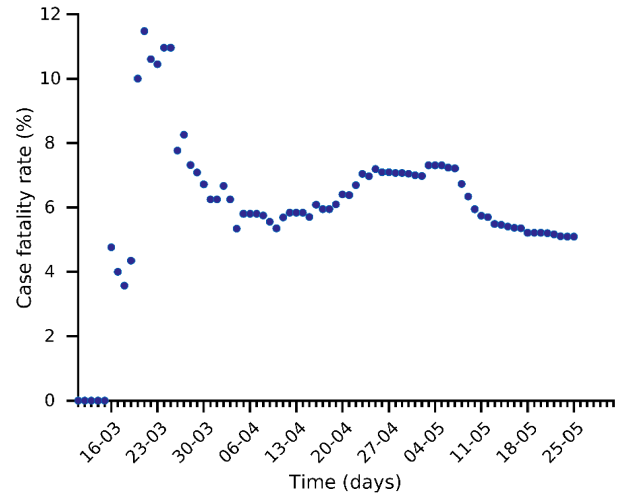




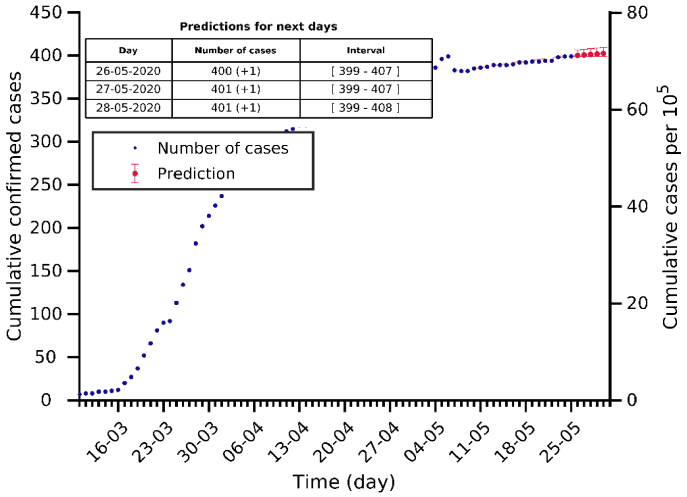


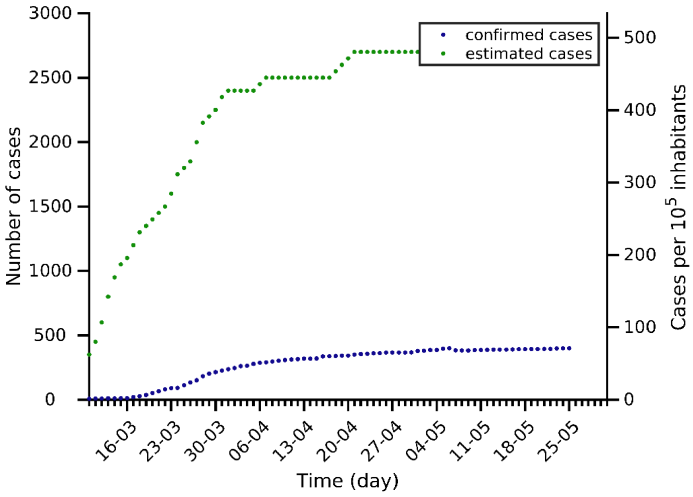


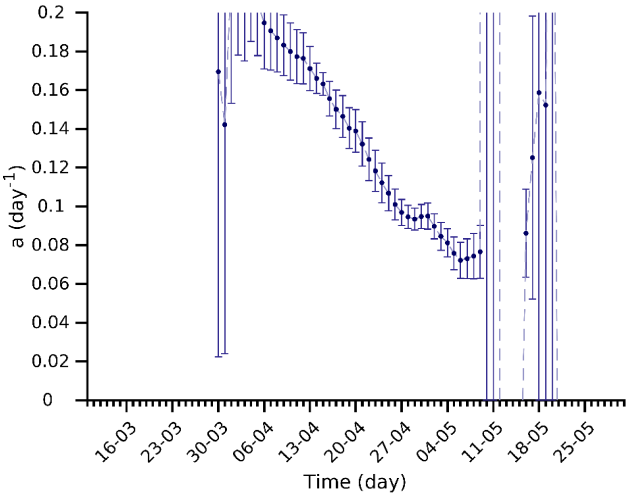


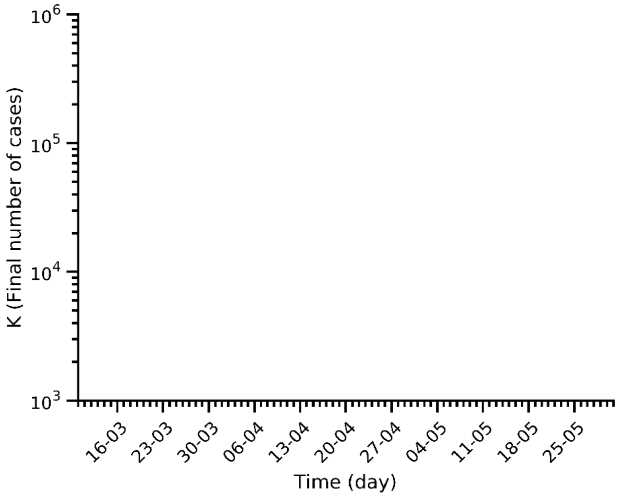


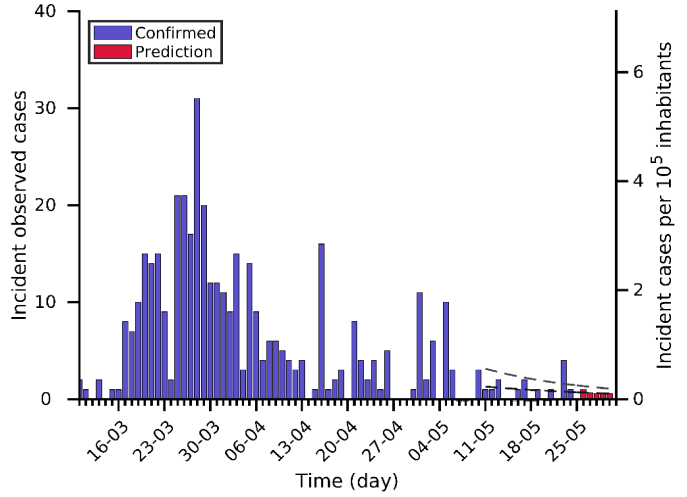
Basilicata 25-05-2020. Population: 0.6M. Current cumulated incidence: 71/10^5

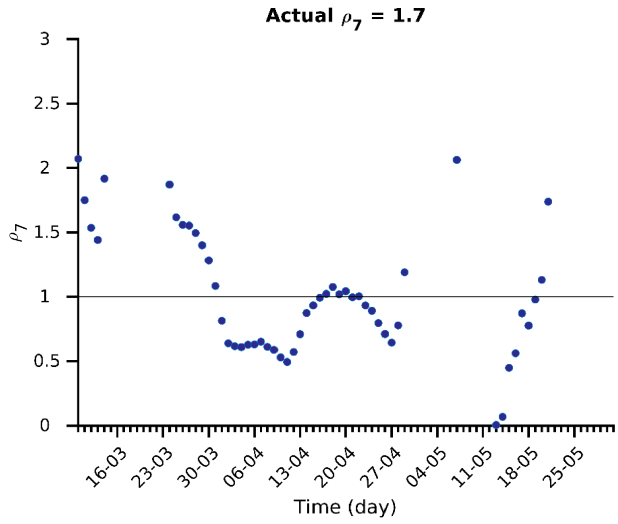


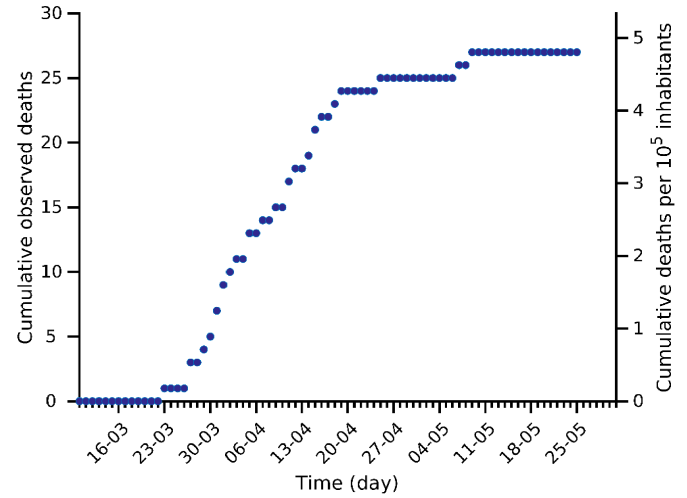


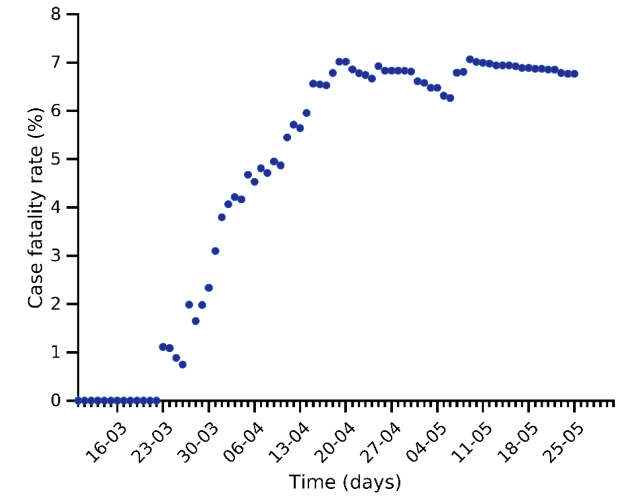












Methods

Methods

(1) Data source

Data are daily obtained from World Health Organization (WHO) surveillance reports[2], from European Centre for Disease Prevention and Control (ECDC)[3] and from Ministerio de Sanidad[4]. These reports are converted into text files that can be processed for subsequent analysis. Daily data comprise, among others: total confirmed cases, total confirmed new cases, total deaths, total new deaths. It must be considered that the report is always providing data from previous day. In the document we use the date at which the datapoint is assumed to belong, i.e., report from 15/03/2020 is giving data from 14/03/2020, the latter being used in the subsequent analysis.

(2) Data processing and plotting

Data are initially processed with Matlab in order to update timeseries, i.e., last datapoints are added to historical sequences. These timeseries are plotted for EU individual countries and for the UE as a whole:

- ✓ Number of cumulated confirmed cases, in blue dots
- ✓ Number of reported new cases
- ✓ Number of cumulated deaths

Then, two indicators are calculated and plotted, too:

- ✓ Number of cumulated deaths divided by the number of cumulated confirmed cases, and reported as a percentage; it is an indirect indicator of the diagnostic level.
- ✓ ρ: this variable is related with the reproduction number, i.e., with the number of new infections caused by a single case. It is evaluated as follows for the day before last report (*t-1*):

$$\rho(t-1) = \frac{N_{new}(t) + N_{new}(t-1) + N_{new}(t-2)}{N_{new}(t-5) + N_{new}(t-6) + N_{new}(t-7)}$$

where $N_{new}(t)$ is the number of new confirmed cases at day *t*.

(3) Classification of countries according to their status in the epidemic cycle

The evolution of confirmed cases shows a biphasic behaviour:

(I) an initial period where most of the cases are imported;
(II) a subsequent period where most of new cases occur because of local transmission.

Once in the stage II, mathematical models can be used to track evolutions and predict tendencies. Focusing on countries that are on stage II, we classify them in three groups:

- Group A: countries that have reported more than 100 cumulated cases for 10 consecutive days or more;
- Group B: countries that have reported more than 100 cumulated cases for 7 to 9 consecutive days;
- Group C: countries that have reported more than 100 cumulated cases for 4 to 6 days.

[2] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
[3] https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases
[4] https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/situacionActual.htm
https://github.com/datadista/datasets/tree/master/COVID%2019 , https://covid19.isciii.es/

(4) Fitting a mathematical model to data

Previous studies have shown that Gompertz model[5] correctly describes the Covid-19 epidemic in all analysed countries. It is an empirical model that starts with an exponential growth but that gradually decreases its specific growth rate. Therefore, it is adequate for describing an epidemic that is characterized by an initial exponential growth but a progressive decrease in spreading velocity provided that appropriate control measures are applied.

Gompertz model is described by the equation:

$$N(t) = K\, e^{-ln\left(\frac{K}{N_0}\right)\cdot e^{-a\cdot(t-t_0)}}$$

where *N(t)* is the cumulated number of confirmed cases at *t* (in days), and N_0 is the number of cumulated cases the day at day t_0. The model has two parameters:

- ✓ *a* is the velocity at which specific spreading rate is slowing down;
- ✓ *K* is the expected final number of cumulated cases at the end of the epidemic.

This model is fitted to reported cumulated cases of the UE and of countries in stage II that accomplish two criteria: 4 or more consecutive days with more than 100 cumulated cases, and at least one datapoint over 200 cases. Day t_0 is chosen as that one at which *N(t)* overpasses 100 cases. If more than 15 datapoints that accomplish the stated criteria are available, only the last 15 points are used. The fitting is done using Matlab's Curve Fitting package with Nonlinear Least Squares method, which also provides confidence intervals of fitted parameters (*a* and *K*) and the R^2 of the fitting. At the initial stages the dynamics is exponential and *K* cannot be correctly evaluated. In fact, at this stage the most relevant parameter is *a*. Fitted curves are incorporated to plots of cumulative reported cases with a dashed line. Once a new fitting is done, two plots are added to the country report:

- ✓ Evolution of fitted *a* with its error bars, i.e., values obtained on the fitting each day that the analysis has been carried out;
- ✓ Evolution of fitted *K* with its error bars, i.e., values obtained on the fitting each day that the analysis has been carried out; if lower error bar indicates a value that is lower than current number of cases, the error bar is truncated.

These plots illustrate the increase in fittings' confidence, as fitted values progressively stabilize around a certain value and error bars get smaller when the number of datapoints increases. In fact, in the case of countries, they are discarded and set as "Not enough data" if $a>0.2\ day^{-1}$, if $K>10^6$ or if the error in K overpasses 10^6.

It is worth to mention that the simplicity of this model and the lack of previous assumptions about the Covid-19 behaviour make it appropriate for universal use, i.e., it can be fitted to any country independently of its socioeconomic context and control strategy. Then, the model is capable of quantifying the observed dynamics in an objective and standard manner and predicting short-term tendencies.

(5) Using the model for predicting short-term tendencies

The model is finally used for a short-term prediction of the evolution of the cumulated number of cases. The predictions increase their reliability with the number of datapoints used in the fitting. Therefore, we consider three levels of prediction, depending on the country:

[5] Madden LV. Quantification of disease progression. Protection Ecology 1980; **2**: 159-176.

- Group A: prediction of expected cumulated cases for the following 3-5 days[6];
- Group B: prediction of expected cumulated cases for the following 2 days;
- Group C: prediction of expected cumulated cases for the following day.

The confidence interval of predictions is assessed with the Matlab function `predint`, with a 99% confidence level. These predictions are shown in the plots as red dots with corresponding error bars, and also gathered in the attached table. For series longer than 9 timepoints, last 3 points are weighted in the fitting so that changes in tendencies are well captured by the model.

(6) Estimating non-diagnosed cases

Lethality of Covid-19 has been estimated at around 1 % for Republic of Korea and the Diamond Princess cruise. Besides, median duration of viral shedding after Covid-19 onset has been estimated at 18.5 days for non-survivors[7] in a retrospective study in Wuhan. These data allow for an estimation of total number of cases, considering that the number of deaths at certain moment should be about 1 % of total cases 18.5 days before. This is valid for estimating cases of countries at stage II, since in stage I the deaths would be mostly due to the incidence at the country from which they were imported. We establish a threshold of 50 reported cases before starting this estimation.

Reported deaths are passed through a moving average filter of 5 points in order to smooth tendencies. Then, the corresponding number of cases is found assuming the 1 % lethality. Finally, these cases are distributed between 18 and 19 days before each one.

[6] At this moment we are testing predictions at 4 days for countries with more than 100 cumulated cases for 13-15 consecutive days, and 5 days for 16 or more days.

[7] Zhou et al., 2020. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet; March 9, doi: 10.1016/S0140-6736(20)30566-3